Hot Carriers in Graphene

Hot Carriers in Graphene

A dissertation presented
by

Justin Chien Wen Song

to
The School of Engineering and Applied Sciences
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Applied Physics

Harvard University
Cambridge, Massachusetts
June 2014

Thesis advisor

Professor Leonid S. Levitov

Author

Justin Chien Wen Song

Hot Carriers in Graphene

Abstract

When energy relaxation between electrons and the lattice is slow, an elevated electronic temperature different from that of the lattice persists. In this regime, hot charge carriers control the energy transport in a material. In this thesis, I show how hot carriers can dominate graphene's response enabling it to exhibit novel properties.

First, I examine how light is converted to electrical currents in graphene and show that hot carriers play an integral role in this multi-stage process. I show that photocurrent in graphene p-n junctions is dominated by a Photo-thermoelectric effect in which a light-induced elevated hot carrier temperature drives a thermoelectric current. Furthermore, I show that the generation and cooling of hot carriers in graphene during photoexcitation proceeds in an unusual way. In the former, carrier-carrier scattering dominates the initial photoexcitation cascade enabling efficient hot carrier generation. In the latter, a new cooling mechanism – disorder-assisted scattering (supercollisions) – dominates electron-lattice cooling over a wide range of temperatures (including room temperature).

Second, I examine the transport characteristics of double layer graphene heterostructures (specifically, G/*h*-BN/G heterostructures). I show that Coulomb coupling results in vertical (out-of-plane) energy transfer between electrons in proximal (but electrically insulated) graphene layers. This couples lateral (in-plane) charge and energy transport of electrons in the two layers to give rise to a new energy-driven Coulomb drag (inter-layer transresistance) that dominates when the two layers are at charge neutrality.

Third, I examine energy transport in charge neutral graphene. I show that the combination of fast carrier-carrier scattering, high electronic quality, and slow electron-lattice cooling (hot carriers) gives rise to a regime of ballistic heat transport. This manifest as electronic energy waves with velocity on the order of graphene's Fermi velocity.

The new phenomena enabled by hot carriers and the ideas/approaches described in this thesis provide a basis with which to exploit hot carrier effects in graphene and opens new vistas for controlling and harnessing energy flows on the nanoscale.

Contents

List of Figures

Acknowledgments

Many people have walked with me and enriched my time here in Cambridge. It takes a village to raise a child, and the community of friends, colleagues, musicians, and mentors have seen me through the most rewarding time of my life yet. My time here has been very blessed because of the imprints these people have left on me.

I owe a deep debt of gratitude to my advisor and mentor, Leonid Levitov. Without him I wouldn't be who or where I am now. He took me in unpolished, and has largely molded my intellectual development. I feel incredibly privileged to have done physics with him, and perhaps more importantly, to have learned how to be a physicist. He taught me by example through the many hours we spent thinking, discussing, and calculating together; this oral tradition has left an indelible mark on me. His clarity of thought has also taught me to write better; I never thought I would learn to write in physics graduate school. In the four years I have spent with him I feel I have matured as a scientist, a writer, and an intellectual. Leonid has been, and remains, a role model for me.

Above all, I am grateful for Leonid's patience, generosity with his time and ideas, and sincere interest in my betterment. The last of which is perhaps a manifestation of his fatherly character; it is this quality, amongst the litany of Leonid's exceptional characteristics, that I have been most impressed with and am privileged to have known him. He has been very thoughtful and intentional in our interactions. In particular, in the last few months he has given me extraordinary support and counsel which has impacted me personally and professionally. I didn't expect it and it came at a time when I needed it most. Thank you for being awesome! Leonya, Spasibo Bolshoe!

Much of my work has been connected with experiments and I have been fortunate to work with Nathan Gabor, Qiong Ma, Pablo-Jarrillo Herrero, Klaas Tielrooij, Frank Koppens, Geliang Yu, and Andre Geim. Nathan and Frank deserve special mention. Nathan was the first experimentalist who believed me; we first met serendipitously in building 13 when I

wandered into his lab. I thank him for the opportunity to dig deep into his data and for playing a big role in getting me excited about science in my G2 year. Frank has been open, and generous with his ideas, data, and time, and I thank him for his confidence in me. I also thank Andrey Shytov, Misha Reizer, and Trung Van Phan for being great theoretical collaborators.

The physicists along the building 4/6C corridor were my comrades. I thank especially my office mate Maksym Serbyn who shared many an espresso, taught me how to integrate, and geeked out over physics. The group members who came before me Mark Rudner, Dima Abanin, Rahul Nandkishore, Nan Gu have been a source of inspiration for me. While we only overlapped briefly, Mark very generously offered his time and counsel. CMT was a very welcoming place to be in, made warmer by: David Mross, Chong Wang, Andrew Potter, Jacob Colbert, Evelyn Tang, Tim Hsieh, Lara Thompson, Rebecca Flint, Karen Micheali, Sam Bieri and more recently, Guy Bunin, Adam Nahum, and Jörn Venderbos who instituted scotch Friday. They all offered their time generously and taught me a lot of what I know. I also thank the building 13 guys: Javier Sanchez-Yamigishi, my gym buddy Valla Fatemi, Andrea Young, Hugh Churchill, Patrick Herring, Spencer Tomarken and Joaquin Rodriguez Nieva. I thank Marc Baldo and Joel Yuen for giving this past year a significant twist. Outside of Cambridge, I thank Brian Skinner and especially Victor Chua, who first showed me the beauty of physics when I was in college.

My homies, Alex Frenzel, Elise Novitski, and Michael Yee, have been a constant source of support and have been my surrogate family here. I thank them for aiding and abetting my hijacking of a G1 desk in Jefferson, prodding me to work with Leonid, supporting my turning to the dark side that is east Cambridge, and perhaps far more important, their consistent faith and confidence in me (even when I wasn't sure of myself). Elena Agapie was the first friend I made in Cambridge and perhaps my biggest fan. I knew right from the

start - when I rode at the back of a Harvard University Police van because of Elena - that we would hit it off. I will always cherish our end of the week dinners, long meandering lunches with Frenzel which kept me honest, mike's shoe and life advice, and Elena's multiple pack and move in a single night challenges. In addition, my homies taught me the value of being genuine and vulnerable, and to always live fully. I love you guys!

Music has always played a huge role in my life and I thank Edward Jones and Christian Lane for the opportunity to sing with the Harvard University Choir, and Tom Brooks, Tibor Krisko, and Bryan Bilyeu at Park Street Church Choir. UChoir was a tremendous experience and I can hardly imagine singing at a higher level. Carols, Compline, the 815am subs, and the fake dimsum Sundays at Yenching crew - Michael Lesley, Kelly Lam, and Felix Boecking - were a mainstay of my first three years. It was also at UChoir that I learned how to tie a bow tie - a significant accomplishment. In my last year, the MIT Gilbert and Sullivan Players became a musical home for me and I thank Barratt Park, Nathaniel Koven, Kathryn Noonan, Laura Plona, and Stephanie Kreutz for all the music, puns, and for Captain Corcoran and Dick Dauntless. I also thank my swing dance friends: Debbie Gaz, Rebecca Graber, and Lauren Hartle for introducing swing to me.

I thank Tout Wang for his thoughtful emails, Gary Sing, Samuel Wong, Travis Stevens for teaching me acceptance, Sean Murphy, Alex McNoughton, Yifeng Wei, Philip Isola, Anna Huang, Amy Shipley, Alex Thomson, Kacey Philips, Jayson Paulose, Anjan Soumyanarayanan, Ilija Zeljkovic, Sean Hart, Justin Wilson, and Amy Koenig for teaching me Latin.

Lastly, I thank my sister, Christine, and my parents for being there for me and being a source of love and encouragement all these years. *Soli Deo Gloria*.

To Christine

How great a being, Lord, is thine, which doth all beings keep!
Thy knowledge is the only line to sound so vast a deep.
Thou art a sea without a shore, a sun without a sphere;
thy time is now and evermore, thy place is everywhere.
John Mason (1683),
"How Shall I Sing That Majesty", The Harvard University Hymn Book

Chapter 1

Introduction

The way in which energy gets distributed amongst a system's available degrees of freedom plays a crucial role in its ability to handle, convert, and utilize energy. Some dramatic examples include the saturation of operating frequencies in integrated electronics due to the large amount of power dissipated in microprocessors [1], and the Shockley-Quiesser limit that sets an upper bound for conversion efficiencies in single-junction solar cells [2]. Two-dimensional materials [3], such as graphene and the atomic layer dichalcogenides, have recently emerged featuring numerous properties with which to manipulate energy flows on the nanoscale. For example, their two-dimensional structure renders electronic states fully exposed, allowing carriers to be extracted via a vertical transfer process (eg. in a sandwich structure). Additionally their high optical activity, absorbing an order of magnitude more sunlight than Silicon layers of similar thickness [4], poise them as ideal optoelectronic materials [5].

Graphene possesses a combination of material properties that make it stand out. These include its broadband absorption [6] allowing it to be responsive to a wide range of frequencies of light, high electronic quality that enables high speed operation [7, 8], gate tunable carrier density that allow ease of varying the size of the Fermi surface and carrier

type (ambi-polarity) [9], and its ease of integration with other materials like Silicon and Boron Nitride through methods such as exfoliation [10].

A key energy attribute of graphene electrons is the ease with which they can be pushed out of thermal equilibrium with the lattice resulting in an elevated electronic temperature different from that of the lattice [11,12]. Importantly, this regime can be characterized by the dynamics of hot (charge) carriers so that electrons control energy transport. As we will see, hot carrier effects in graphene exist in a wide range of technologically relevant temperatures including room temperature. As a result, graphene is an attractive material for high speed energy transduction.

A sustained population of hot carriers stems from the inefficiency of electron-lattice energy relaxation in graphene. Strong carbon-carbon bonds that give graphene's lattice its rigidity also result in a high optical phonon frequency, $\omega_0 = 200\,\text{meV}$. This large value of the optical phonon energy renders optical phonon scattering inefficient below a few hundred kelvin. At the same time, the already weak scattering between electrons and long-wavelength acoustic phonons is further constrained by the large mismatch in Fermi velocity and sound velocity $v/s \approx 100$[1]. As a result, once the electrons are heated up they stay out of thermal equilibrium with the lattice over long times and extended spatial lengthscales. Hot carriers have been studied in a variety of other systems including semiconductors like GaAs [13, 14] and metals [15]. In these other materials, hot carriers only exist at very low temperatures or under intense pumping. In contrast, hot carriers can exist in graphene even at room temperature and under weak driving.

While the proliferation of hot carriers in graphene impacts its energy transport which is *charge neutral*, experimental and technologically relevant responses are typically

[1] As discussed later in the introduction, this constrains the cooling process to the emission of only very long wavelength acoustic phonons to make this cooling process very inefficient.

most sensitive to *charged* modes. A further ingredient - coupling between charge neutral energy modes and charged modes - is required for hot carriers in graphene to have a significant impact on its response properties. This is provided by strong thermoelectricity in graphene [16,17] that allow charged modes to be strongly coupled to energy modes: a spatial temperature gradient can generate a voltage drop (Thermopower effect), and a charge current can drive heat flow (Peltier effect). Combined with a sustained hot carrier distribution, new kinds phenomena can manifest in graphene's "hot carrier" regime.

In this thesis, I will discuss the generation and dynamics of hot carriers, and importantly, how they can lead to a range of new phenomena in graphene. As will become clear, a unique combination of material properties make graphene an ideal venue to realize hot carrier effects. I will argue that hot carriers drive many of graphene's responses (eg. graphene's response to light is dominated by hot carrier effects), and qualify it as an exciting new material in which to mold the flow of energy. In the following, I lay out the main arc of this thesis and discuss the larger context in which this work resides in.

A particularly instructive setting in which to examine the impact of hot carriers is graphene's photoresponse: how does graphene respond to light impinging on it? One way of describing light-matter interaction in graphene is as follows: Light impinging on graphene injects energy into graphene's electrons and pushes the electronic system out of thermal equilibrium with the lattice; light is used as a means of heating graphene electrons. This way of thinking about optoelectronics in graphene stands in stark contrast to the lens conventionally used to view photoresponse: the Photovoltaic effect. In the Photovoltaic effect, light creates electron-hole pairs that are subsequently separated by the bulit-in field at a p-n junction to create a photocurrent [18]. Indeed, prior to work described in this thesis, photocurrent in graphene p-n junctions was primarily attributed to the PV effect [5, 7, 19]. In the first half of this thesis, I will address how graphene responds to light by focusing on

energy flows in graphene.

In chapter 2, I will argue that hot carrier effects dominate graphene's photocurrent response, and discuss methods to distinguish it from the Photovoltaic effect. This prediction [20] and subsequent experimental confirmation [21] that hot carrier effects are responsible for the intrinsic photocurrent response of graphene has prompted a fresh look at what limits photocurrent, and how to exploit this new mechanism for converting light energy to electrical current in graphene. For instance in this new "hot carrier" regime, hot carrier temperature, and electron-lattice cooling times in graphene, replaces more traditional photovoltaic concepts like total number of electron-hole pairs generated, and electron-hole recombination times. Indeed, engineering both device length scales to be smaller than electron-lattice cooling lengths [22,23] and thermal environment [24] have since been used to create sensitive graphene based bolometers. The work described in chapter 2 was performed just as there was an upsurge in technology interest surrounding graphene photodetectors [5], in part because of graphene's broadband absorption of light [6] and possibility for fast photodetection [7]. This provided a microscopic basis for exploiting hot carriers in graphene based photodetection.

One pertinent question that arises given the role hot carriers play in graphene's photoresponse is what microscopic mechanisms control the generation and cooling of hot carriers after initial photoexcitation. In the former, fast scattering processes (tens of fs) such as carrier-carrier scattering with the ambient carrier distribution generate hot carriers [25]. Equivalently, these fast processes determine how much energy is captured by the electronic system (manifested as an elevated electronic temperature) when energy is pumped into graphene. Once a hot carrier distribution is established, slower cooling processes (up to hundreds of ps), such as phonon emission, determine how long the carriers stay hot [11,12,26]. The separation of time scales between the generation and cooling of hot carriers allow us to

consider these processes separately. Control over both generation and cooling processes will provide means to manipulate the energy flows in graphene, key in exploiting it as a future energy material.

First, I will lay out the microscopic processes that control hot carrier cooling in Chapter 3. I will show how a new electron-lattice cooling mechanism, unknown prior to work described in this thesis, dominates the cooling of hot carriers in a large technologically relevant range, from tens of Kelvin to room temperature. This theoretical prediction triggered a number of experimental efforts which accurately measured the electron-lattice cooling dynamics [27], and electron-lattice cooling power [28] experimentally confirming that this new electron-lattice cooling mechanism dominates cooling in graphene.

The technological promise of graphene energy harvesting has prompted huge interest in Auger processes in graphene [25] because these could in principle allow for significant efficiencies to be gained [20]. Photoexcitation has emerged as an attractive way of generating hot carriers in graphene [29, 30]. In this approach, high energy photoexcited electron-hole pairs can heat up the ambient carriers in graphene via carrier-carrier scattering (Auger-type processes) [29, 30]; energy from the primary photoexcited carriers is captured by graphene's ambient carriers to give a hotter electron temperature (hot carriers). Indeed, high electronic temperatures in graphene have been observed under intense irradiation [31]. Graphene's large interaction parameter, $\alpha \approx 2.2$ (in vacuum) [32], gapless spectrum, and tunable carrier density make it an ideal venue to analyze the role of Auger type processes in the energy relaxation of high energy photoexcited carriers (photoexcitation cascade).

In Chapter 4, I describe how the relaxation of high-energy photoexcited carriers proceeds giving special attention to the competition between Auger-type and Phonon emission processes. In a combined theory-experiment effort, the branching ratio between these two process types was extracted favoring Auger type processes [30]. The large fraction of

energy captured by the electronic system has triggered renewed interest in the so-called "hot carrier solar cell" [33] in which a hot carrier distribution is used to drive a nanoscale heat engine [34]. Hot carrier solar cells circumvent the Shockley-Quiesser limit promising to achieve high light-to-electricity energy conversion efficiencies [34].

In Chapter 5, I describe a framework for spatial patterns of photocurrent response in gapless materials, such as graphene. While the theory presented is not specific to hot carrier assisted photoresponse, it provides the necessary link to understanding how local photoexcitations lead to currents in far-away contacts. It mirrors the famous Shockley-Ramo theorem [35, 36] for currents induced in contacts by moving charges in vacuum tubes and semiconductors. This framework for gapless materials conveniently explains several new features of photocurrent seen in recently discovered gapless materials [37, 38].

One of the attractions of the recently discovered Van der Waals' materials, that include graphene and atomic layer dichalcogenides, is the ease with which vertical heterostructures can be made out of graphene and other two-dimensional materials [3]. This ability to engineer in the vertical direction (third dimension) provides a direct way to tune the electronic environment in which the two-dimensional electron gas in graphene (or other two dimensional material) resides in. This has led to the observation of new electronic phenomena like the Hofstadter Butterfly [39–41], replica Dirac points [42, 43], amongst others.

This new capability also opens the possibility of studying new pathways of energy transport in which electron interactions play a central role. An important example of such a vertical structure is a vertical stack Graphene/Boron Nitride/Graphene; Boron Nitride is an insulator with a large gap that effectively blocks vertical particle tunneling between the two graphene layers. The small layer separations ($d \approx$1-2 nm) and the weak electron-lattice coupling embodied by hot carrier populations that are sustained over long times, open up the possibility for electron interactions to play a dominant role in energy flows.

Can inter-layer carrier-carrier scattering mediate energy transfer between graphene layers? In Chapter 6, I discuss how fast inter-layer carrier-carrier scattering mediates fast energy transfer between the electron systems in the two layers. Strikingly, vertical energy transfer mediates a trans-resistance, called drag resistance, between the two layers (without particle exchange), giving rise to a new mechanism for drag resistance - energy-driven drag - that dominates the drag response close to charge neutrality. Indeed, recent measurements [44] find drag resistance consistent with a dominant energy-driven drag. Similar to photoresponse discussed above, the dominant role of energy-driven drag can be traced back to hot carriers in graphene that can be sustained over long length scales.

Drag resistance has a long history as a sensitive probe of electron interactions in two-dimensional electron gases [45, 46]. The ease with which double layer graphene heterostructures with $d < \lambda$ can be made allows the "strong coupling" regime to be accessed. Here λ denotes a characteristic electron length scale such as the characteristic electron length scales such as the de Broglie wavelength and screening length. This is the regime where inter- and intra-layer electron-electron interactions are equally strong giving drag which can be large. At the time that the work described in Chapter 6 (and 7) was performed, there was interest in using graphene drag measurements to probe interlayer correlations [44]. [2] Our work, performed in parallel to the experimental work, provided a clear mechanism in which to understand the anomalous results that the experimentalists were seeing [44].

In Chapter 7, I discuss the fate of energy-driven drag when a perpendicular magnetic field is applied and compare it with other mechanisms for drag resistance. While energy transfer between layers does not change, qualitatively new features in drag resistance (trans-

[2]For instance, one tantalizing possibility was that double layer graphene heterostructures might host exciton condensates. The anomalous signatures seen in the experiment [44] particularly when magnetic field was applied created hope in the field that such exciton condensates were responsible for the new features observed close to charge neutrality. Our work was performed in part to see if other mechanisms could be responsible for the anomalous drag signatures.

port) appear. Recent measurements [44] have confirmed several of the qualitative features predicted. The study of thermal effects in hall-type measurements have been of interest to the mesoscopic community. This is because, when complemented with more conventional charge hall measurements, they reveal a fuller picture of how particles are transported in exotic states that can occur in a (strong) magnetic field [47, 48]. Magneto and Hall energy-driven drag in graphene double layers provides a further means in which to probe energy transport in a magnetic field.

As illustrated in Chapter 6 & 7, electron interactions can mediate energy flow and can even dominate its charge transport characteristics. Can electron interactions give rise to qualitatively new kinds of energy transport? In chapter 8, I discuss how a ballistic wave-like energy propagation can occur in single layer graphene. This collective mode of energy transport occurs in the "hydrodynamic regime" wherein carrier-carrier scattering, occurs far faster than either momentum relaxation to impurities in graphene, or energy relaxation to its lattice. While this new energy transport mechanism has yet to be realized experimentally, it provides a tantalizing possibility of achieving high speed energy transduction given that energy waves in graphene travel at anomalously high speeds, $\approx 10^6\,\mathrm{cms}^{-1}$.

The rest of the introduction is structured as follows: I will provide a brief overview of electrons and phonons in graphene and their interaction. Finally, I provide a brief summary of the main results of this thesis.

1.1 Electrons and Phonons in Graphene

Graphene consists of an atomically thin two-dimensional honeycomb structure of carbon atoms (see Fig. 1.1). The structure of carbon atoms in graphene determine how both the electrons and phonons in graphene behave and interact with each other. Below I review

some of their basic characteristics.

Graphene's structure arises from chemistry between adjacent carbon atoms determined by the four valence electrons each carbon atom has. The atomic orbitals s, p_x, and p_y orbitals of adjacent carbon atoms hybridize into three sp^2 orbitals, forming strong $\sigma-$bonds in the plane. This gives graphene's hexagonal lattice its rigidity. The remaining p_z lies normal to the plane. The overlap of p_z orbitals on adjacent carbon sites [3] gives rise to the low-energy behavior of graphene electrons.

1.1.1 Dirac Electrons

The honeycomb structure of graphene's lattice gives rise to the electron's unique characteristics (Fig. 1.1), and can be understood from considering electrons hopping between adjacent sites in the p_z orbitals. This hopping is amenable to a simple tight binding description of the low energy behavior of graphene electrons

$$H_0 = \sum_{\langle ij \rangle} t\psi_i^\dagger \psi_j \tag{1.1}$$

where $\psi_i^\dagger$ creates an electron on site i and ψ_j annihilates an electron on site j, and $t \approx 3\,\mathrm{eV}$ is the nearest neighbour hopping energy, $\langle ij \rangle$ indicates hopping only to nearest neighbour sites. Graphene's honeycomb structure can be tessellated by a triangular Bravais lattice with a unit cell consisting of two carbon atoms (one called A and the other called B, see Fig. 1.1a). As a result, the tight-binding Hamiltonian above can be conveniently diagonalized in momentum space, taking the form of a 2×2 matrix $H = \sum_{\mathbf{k}} c^\dagger(\mathbf{k}) H_0(\mathbf{k}) c(\mathbf{k})$, where $\mathbf{k}$ is the wavevector [4], and $c(\mathbf{k}) = (c_{\mathbf{k}}^A, c_{\mathbf{k}}^B)$ has two components that arise from the two carbon

[3] so called π-bonding

[4] momentum, $\hbar\mathbf{k}$, in a crystal lattice is well defined up to a reciprocal lattice vector. For the hexagonal structure pictured in Fig. 1.1a, these are

$$\mathbf{b}_1 = 2\pi/3a(1, \sqrt{3}), \quad \mathbf{b}_2 = 2\pi/3a(1, -\sqrt{3}) \tag{1.2}$$

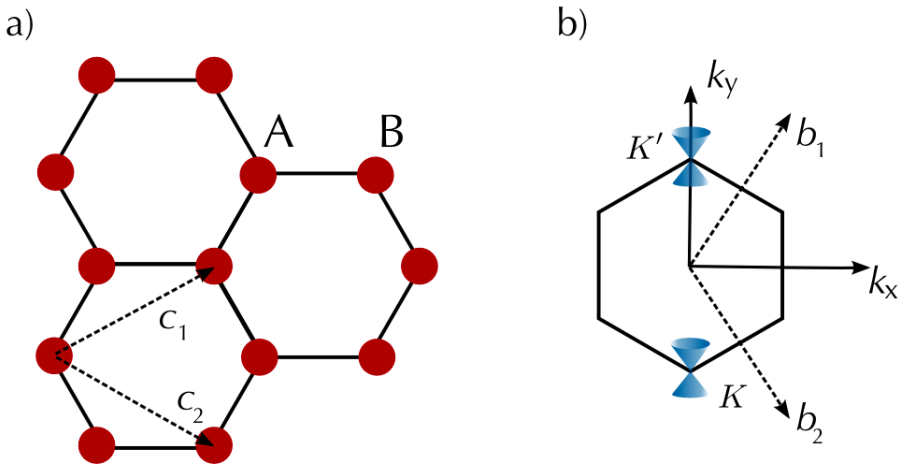

Figure 1.1: Graphene's honeycomb lattice and its Brillouin zone. a) Lattice structure of graphene is tesselated by a triangular lattice (lattice vectors $\mathbf{c}_1$ and $\mathbf{c}_2$) with a 2-atom unit cell labeled A and B. b) First Brillouin zone of graphene with reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$. The energy spectrum of graphene features a gapless Dirac cones located at $\mathbf{K}$ and $\mathbf{K}'$ points.

atoms in each unit cell (called a spinor). This yields

$$H_0 = \begin{pmatrix} 0 & \tau(\mathbf{k}) \\ \tau^*(\mathbf{k}) & 0 \end{pmatrix}, \quad \tau(\mathbf{k}) = t\sum_j^3 e^{i\mathbf{k}\cdot\delta_j} \tag{1.3}$$

where the off-diagonal structure comes from the hopping of A site to nearest neighbour B sites and vice versa, and δ_j are the vectors in real space linking nearest neighbours. This hamiltonian has eigenvalues that is symmetric about $\epsilon = 0$ [49], and importantly is gapless.

Pristine graphene, called undoped graphene [5], fills half the electronic states since each carbon atom contributes one electron to the p_z orbital system. As a result of particle hole symmetry (see above), the chemical potential for undoped graphene is at $\epsilon = 0$; this is also called half-filling. While Eq. 1.3 gives rise to bandstructure throughout the Brillouin zone, for the purposes of this thesis we will concentrate on the low-energy behavior close to

where $a \approx 0.14\,\text{nm}$ is the inter-carbon atom distance.

[5] since it is charge neutral

$\epsilon = 0$ since technologically relevant graphene devices have chemical potentials in the vicinity of $\epsilon = 0$. Chemical potential at half filling, $\mu = 0$, is only realized in the pristine case.

The energy spectrum close to $\epsilon = 0$ is concentrated at two inequivalent points in the Brillouin zone, $\mathbf{K} = -4\pi/(3\sqrt{3}a)\hat{\mathbf{y}}$, and $\mathbf{K'} = 4\pi/(3\sqrt{3}a)\hat{\mathbf{y}}$; at K and K' the energy is zero. In the vicinity of $\mathbf{K}$ and $\mathbf{K'}$, and writing $\mathbf{k} = \mathbf{K}+\mathbf{p}$, $\mathbf{k} = \mathbf{K'}+\mathbf{p}$ yields $\tau(\mathbf{k} = \mathbf{K'}+\mathbf{p}) \approx 3ta/2(ip_x - p_y)$, and $\tau(\mathbf{k} = \mathbf{K} + \mathbf{p}) \approx 3ta/2(ip_x + p_y)$. These produce the hamiltonian

$$H_{\mathbf{K}} = v\hbar\sigma \cdot \mathbf{p}, \quad H_{K'} = v\hbar\tilde{\sigma} \cdot \mathbf{p}, \quad \sigma = (\sigma_x, \sigma_y), \quad \tilde{\sigma} = (-\sigma_x, \sigma_y) \tag{1.4}$$

where $v = 3ta/2 \approx 10^8 \mathrm{cm\,s^{-1}}$, and $\mathbf{p}$ has been rotated to yield the form more conventionally used. Eq. 1.4 is called the Dirac hamiltonian as it closely resembles the Dirac equation in quantum electrodynamics but with v and sublattice occupation/pseudospin replacing the speed of light and real spin; in the same spirit, $\mathbf{K}, \mathbf{K'}$ are called Dirac points. The linear approximation of the spectrum (above) approximates the complete bandstructure of graphene very well for a wide range of energies and only starts to deviate for $\epsilon \approx \pm 1\,\mathrm{eV}$.

Eq. 1.4 only describes the single particle physics of graphene electrons, and a complete pictures of electronic behavior requires electron-electron interactions. This is provided by long ranged Coulomb interactions between electrons which can be described by

$$\mathcal{H}_{\mathrm{el-el}} = \frac{1}{2} \sum_{\mathbf{q},\mathbf{k},\mathbf{k'},i,j} V(\mathbf{q}) \psi^{\dagger}_{\mathbf{k}+\mathbf{q},i} \psi^{\dagger}_{\mathbf{k'}-\mathbf{q},j} \psi_{\mathbf{k'},j} \psi_{\mathbf{k},i} \tag{1.5}$$

where $V(\mathbf{q}) = 2\pi e^2/|\mathbf{q}|$ is the bare Coulomb interaction in momentum space (two-dimensional). Finite charge density can also screen this Coulomb interaction and modifies $V(\mathbf{q})$. While electron-electron interactions can lead to some renormalization of parameters (such as velocity) in graphene, they are not expected to change the qualitative character of the electronic ground state in graphene [32]; the Fermi liquid is a good ground state. However, as I detail in Chapter 4, 6-8, electron interactions through carrier-carrier scattering and the energy exchanged in these events play a critical role in the transport and flow of energy in graphene

and graphene heterostructures. Indeed, the gapless spectrum and relatively high density of carriers in doped graphene grant a large phase space for electron-electron scattering. In this thesis, the primary role of electron-electron interactions is to provide a fast scattering channel that mediates the flow of energy.

1.1.2 Phonons and Electron-Phonon Coupling

Phonons are lattice vibrations of the crystal, and there are two principal types of phonons in graphene: Optical[6] and Acoustic. They can be described by a phonon hamiltonian $H_{\text{ph}} = \sum_{\mathbf{q},i} \omega_{\mathbf{q},i} b^\dagger_{\mathbf{q},i} b_{\mathbf{q},i}$, where $\omega_{\mathbf{q},i}$ is the spectrum of the phonons, and $b^\dagger_{\mathbf{q},i} b_{\mathbf{q},i}$ are phonon creation and annihilation operators; the in-plane optical and acoustic phonon spectra can be approximated as

$$\text{Optical phonons} : \omega_0 \approx 200\,\text{meV}, \quad \text{Acoustic phonons} : \omega_{\mathbf{q}} = s|\mathbf{q}|, \quad s \approx 10^6 \text{cm}\,\text{s}^{-1} \tag{1.6}$$

respectively. The high optical phonon energy arises from the stiff carbon-carbon bonds that make up graphene's structure. Out-of-plane (called flexural) phonons can also occur. However, graphene devices are conventionally found on substrates and the flexural phonon contribution to transport and energy flow characteristics are expected to be small since they can get pinned, gapped, and stiffened by contact with the substrate. They only become important in suspended samples of graphene and we will briefly consider them in Chapter 3.

The electron density, $n(\mathbf{r})$, sees the lattice of ions (carbon atoms in graphene). Indeed their interaction energy can be modeled as $V_{\text{el}-\text{ion}} = \int d^2\mathbf{r} e n(\mathbf{r}) \sum_j V_{\text{ion}}(\mathbf{r} - \mathbf{R}_j)$, where V_{ion} is the potential of each individual carbon atom, and $\mathbf{R}_j$ is the position of the carbon atoms. Vibrations of the carbon atoms in phonon modes displace the carbon atoms

[6]Optical phonons occur in systems with one than one atom in each unit cell

from their equilibrium positions, $\mathbf{u}_j = \mathbf{R}_j - \mathbf{R}_j^0$. Here $\mathbf{u}_j$ is the displacement vector of a carbon atom and $\mathbf{R}_j^0$ is the equilibrium position of the carbon atoms in the lattice. As a result, $V_{\mathrm{el-ion}}$ changes from its equilibrium value; phonons can interact with the electron density[7]. For acoustic phonons in graphene, the electron-phonon interaction can be modeled in the usual way [11, 12]

$$H_{\mathrm{el-ph}} = \sum_{\mathbf{q}} g\sqrt{\omega_{\mathbf{q}}}\left(b_{\mathbf{q}} + b^{\dagger}_{-\mathbf{q}}\right)n_{\mathbf{q}}, \quad g = D/\sqrt{2\rho s^2} \tag{1.7}$$

where $D \sim 20\mathrm{eV}$ [11, 12] is the deformation potential constant and ρ is the mass density of the graphene sheet. We note that translational invariance of the lattice makes coupling between electrons and long-wavelength acoustic phonons weak.

The inefficiency of scattering between electrons and phonons in graphene allows for a sustained population of hot carriers. To see this more concretely, we can consider the scattering rate between electrons with single phonons (acoustic or optical)

$$W_{\mathbf{k}',\mathbf{k}} = \frac{2\pi}{\hbar}\sum_{\mathbf{q}} F_{\mathbf{k},\mathbf{k}'}|M|^2\left[\delta_{\mathbf{k}'-\mathbf{k}-\mathbf{q}}N_{\omega_{\mathbf{q}}}\delta_{+} + \delta_{\mathbf{k}'-\mathbf{k}-\mathbf{q}}(N_{\omega_{\mathbf{q}}}+1)\delta_{-}\right] \tag{1.8}$$

where $\delta_{\pm} = \delta(\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}} \mp \omega_{\mathbf{q}})$, $\mathbf{q}$ is phonon momentum and $N_{\omega_{\mathbf{q}}} = 1/(e^{\beta\omega_{\mathbf{q}}} - 1)$ is the Bose distribution, and $F_{\mathbf{k},\mathbf{k}'}$ is the graphene coherence factor. Here $M = g\sqrt{\omega_{\mathbf{q}}}$ for acoustic phonons and $M = g_0$ for optical phonons. Here $g_0 = \frac{2\hbar^2 v}{\sqrt{\rho\omega_0 a^4}}$ [50]. For acoustic phonon scattering, simultaneous momentum and energy conservation (embodied in the delta functions of Eq. 1.8) mean that only phonons with energy at most $\omega = 2s\hbar|\mathbf{k}|$ can be exchanged. Here $\mathbf{k}$ is the initial wavevector of the incoming electron. For electrons around the Fermi surface (in a hot Fermi distribution), this translates to $\omega = 2s\hbar k_F = 2T_{\mathrm{BG}} \approx 10K$ acoustic phonons exchanged (for typical doping). As a result, only very long wavelength (or low energy) acoustic phonons are emitted in the cooling process, making this process very ineffective.

[7] Noting that $\mathbf{u}_{\mathbf{q}} = \ell_{\mathbf{q}}\left(b^{\dagger}_{\mathbf{q}} + b_{\mathbf{q}}\right)$ are connected with phonon modes, and expanding $V_{\mathrm{el-ion}}$ in $\mathbf{u}$ allows one to obtain the electron-phonon interaction, $H_{\mathrm{el-ph}}$. Here $\ell_{\mathbf{q}} = \sqrt{\hbar/2M\omega_q}$ and M is the mass of the ion.

To give a sense of how ineffective this process is, it is useful to examine the electron-lattice cooling rate from scattering off single acoustic phonons defined as $dT_{\rm el}/dt = -\gamma(T_{\rm el} - T_0)$, $T_{\rm el} \approx T_0$ [8]. Here T_0 is the lattice temperature and $T_{\rm el}$ is the electron temperature. These give slow cooling rates [11, 12]

$$\gamma = \frac{3D^2\mu^3}{4\pi^2\hbar^3\rho v^4 k_{\rm B} T_{\rm el}} \approx 0.87 \frac{(\mu\,[\text{meV}]/100)^3}{T_{\rm el}\,[\text{K}]/300}\text{ns}^{-1}, \tag{1.9}$$

where $D = 20\,\text{eV}$, $\rho = 7.6\times10^{-8}\,\text{g/cm}^2$, and $v = 10^6\,\text{m/s}$ are the deformation potential, mass density, and Fermi velocity for graphene. For realistic values $\mu = 100\,\text{meV}$ and $T_{\rm el} \sim 300K$, this yields timescales as long as a few nanoseconds.

Similarly, the cooling power $\mathcal{J} = \sum_{\mathbf{k},\mathbf{k}'}(\epsilon_k - \epsilon_{k'})W_{\mathbf{k}',\mathbf{k}} f(\epsilon_k)(1 - f(\epsilon_{k'}))$ from optical phonons is also inefficient. This stems from the large optical phonon energy, $\omega_0 \approx 200\text{meV}$. Energy conservation in Eq. 1.8 means that $\epsilon_{k'} = \epsilon_k - \omega_0$. As a result, even though a large energy can be carried away by optical phonons, the product of occupation factors $f(\epsilon_k)(1 - f(\epsilon_k - \omega_0))$ suppresses this rate for temperatures smaller than $\omega_0 \approx 2500\,\text{K}$. Importantly, this renders optical phonons ineffective for cooling in the range of technologically important temperatures (tens of Kelvin to room temperature). A full discussion of electron-lattice cooling mechanisms in graphene is undertaken in Chapter 3.

These long time scales for cooling between electrons and the lattice in graphene even at room temperature set the stage for hot carrier effects to play a key role in graphene's response.

[8] γ is obtained by noting that for $T_{\rm el} \approx T_0$, linearizing the cooling power $\mathcal{J} = \sum_{\mathbf{k},\mathbf{k}'}(\epsilon_k - \epsilon_{k'})W_{\mathbf{k}',\mathbf{k}} f(\epsilon_k)(1 - f(\epsilon_{k'}))$ allows us to express $\mathcal{J} = \gamma C_{\rm el}(T_{\rm el} - T_0)$. Here, $C_{\rm el}$ is the electron heat capacity.

1.2 Main Results of Thesis

In this thesis, I discuss how energy transport through the dynamics of hot carriers can lead to a range of new phenomena in graphene that include its photoresponse (Chapter 2-5), charge transport characteristics (Chapter 6 & 7), and energy transport (Chapter 8). Additionally, since the generation and cooling of hot carriers can play a determining role in the magnitude hot carrier effects, I will discuss the microscopics of the generation (Chapter 4) and cooling of hot carriers (Chapter 3).

1.2.1 Hot Carriers and Graphene's Intrinsic Photocurrent Response

The electron-lattice cooling bottleneck in graphene allows hot carrier temperatures to be sustained over long lengthscales. In Chapter 2, I will describe how hot carriers dominate the photocurrent generation in p-n junction, so that a light-induced elevated hot carrier temperature drives a thermoelectric current: the Photothermoelectric effect. Additionally, I predict key qualitative signatures of the hot carrier dominated photocurrent response in graphene, the most striking of which is multiple sign changes of photocurrent as the polarity of carriers on either side of a p-n junction are varied.

In the second part of Chapter 2, I detail a theory-experiment collaboration which confirmed the signatures of the hot carrier dominated response in detail, including multiple sign changes that manifest as a six-fold photovoltage pattern, and a hot carrier temperature that was tunable by gate.

This chapter provides a broad overview of the multi-stage process of conversion of light into electrical current and serves as an illustrative example of how hot carriers can affect graphene's response. It also shows where hot carrier cooling mechanisms (discussed in Chapter 3) and hot carrier generation mechanisms (discussed in Chapter 4) enter into the photocurrent response signal.

1.2.2 Supercollision Electron-Phonon Scattering and Cooling

In chapter 3, I predict that graphene is a unique system where disorder-assisted scattering (supercollisions) dominates electron-lattice cooling over a wide range of temperatures, up to room temperature. This is so because for momentum-conserving electron-phonon scattering, the energy transfer per collision is severely constrained due to a small Fermi surface size, allowing only acoustic phonons with energy $k_B T_{\mathrm{BG}} = s\hbar k_F$ to be emitted. In contrast, the higher order (supercollision) process allows thermal phonons with energy $k_B T$ to be emitted vastly increasing energy relaxation. The characteristic T^3 temperature dependence and power-law cooling dynamics provide clear experimental signatures of this new cooling mechanism. The cooling rate can be changed by orders of magnitude by varying the amount of disorder providing a control knob for a variety of new applications that rely on hot-carrier transport.

I also briefly discuss the experimental confirmation [27, 28, 51] of this new cooling mechanism, and detail the variety of supercollision-type processes that include scattering off ripples, two-phonon processes, and the emission two flexural phonons. The last two are expected to dominate in pristine graphene, where disorder is eliminated.

1.2.3 Hot Carrier Generation and Photoexcitation Casacde

The conversion of light into free electron-hole pairs constitutes the key process in the fields of photodetection and energy harvesting. In materials with strong electron-electron interactions, photoexcitation can trigger a cascade in which multiple particle-hole excitations are generated. In Chapter 4, I analyze the cascade originating from intraband carrier-carrier scattering processes (in this thesis, we call this type of Auger process impact-excitation) in graphene in which many hot carriers are generated by a single absorbed photon. Comparing this with phonon emission, I argue that the impact-excitation process dominates the energy

relaxation of high energy photoexcited carriers allowing them to relax primarily by directly heating up the electronic system; hot carriers are generated efficiently via photoexcitation. I show that the number of generated carriers has a strong dependence on doping (gate tunability). Linear scaling with photon energy is predicted for the number of pairs and for the duration of the cascade. These dependencies, along with a sharply peaked angular distribution of excited carriers, provide clear experimental signatures of a carrier-carrier scattering dominated photoexcitation cascade.

I also discuss how hot carriers gives rise to a clear picture of photoexcitation from initial absorption of photon to final relaxation back to equilibrium. Additionally, I discuss how the hot carrier temperature can be extracted experimentally using the electronic temperature dependence of terahertz conductivity. In an experiment-theory collaboration, a pump-probe study tracked the photoexcitation cascade using pump-induced changes in THz photoconductivity to extract the efficiency of impact-excitation process vs optical phonon emission, concluding that the impact-excitation process was several times more efficient.

1.2.4 Long-Range Photocurrent Response in Gapless Materials

In chapter 5, I detail a general Shockley-Ramo-type framework that can be used to describe long-range spatial patterns of photocurrent response in gapless materials. This approach helps to understand the striking features of the observed patterns, such as the directional effect and the global character of photoresponse. I illustrate this approach by examining specific examples, and show that the photoresponse patterns can serve as a powerful tool to extract information about symmetry breaking, inhomogeneity, chirality, and other local characteristics of the system. Short response times, originating from the nonlocality, make photocurrent uniquely sensitive to charge dynamics.

1.2.5 Vertical Energy Transfer and Energy-driven Drag in Graphene

Hot carriers can also affect the transport characteristics in graphene devices. The most striking of which occurs in Coulomb drag of double layer graphene heterostructures (specifically, G/h-BN/G stacks).

In chapter 6, I describe how Coulomb coupling in graphene heterostructures results in vertical energy transfer between electrons in proximal layers. In the presence of correlated density inhomogeneity in the layers, vertical energy transfer has a strong impact on lateral charge transport; in this chapter, I focus on the $B = 0$ case. In particular, for Coulomb drag, its contribution dominates over conventional momentum drag near zero doping. The dependence on doping and temperature, which is different for the two drag mechanisms, can be used to separate these mechanisms in experiment. I detail distinct features such as a peak at zero doping and multiple sign reversal, which provide diagnostics for this new drag mechanism.

In chapter 7, I describe how this energy-driven mechanism gives rise to magneto and Hall drag for $B \neq 0$. I argue that even in the absence of density inhomogeneity (required in the previous chapter for $B = 0$ energy drag), charged and energy modes become strongly coupled when a B field is applied. This regime is characterized by strong magnetodrag and Hall drag between two graphene sheets, originating from long-range energy currents and spatial temperature gradients. The energy-driven effects arise in a wide temperature range, feature an unusually strong dependence on field and carrier density with giant values of magnetodrag and Hall drag occuring at classically weak fields. Under realistic conditions, energy transport dominates in a wide temperature range, giving rise to a universal value of energy driven drag which is essentially independent of the electron-electron interaction. In the second part of chapter 7, I also compare and contrast momentum and energy-driven mechanisms on an equal footing using a two-fluid model.

Many of the features of the above mechanism for Coulomb drag including the sign of drag at charge neutrality for both $B = 0$ and $B \neq 0$ and the giant values for magento and Hall drag attained in a magnetic field have been since seen in experiment [44].

1.2.6 Collective Wave-like Energy Transport

Lastly in chapter 8, I detail how a ballistic energy transfer mode, with heat propagation governed by a wave equation rather than a diffusion equation, can be realized for a thermal electron-hole plasma in graphene. This new behavior originates from rapid exchange of energy and momentum in particle collisions leading to energy propagation as a collective weakly-damped oscillation; it occurs in the hydrodynamic regime where the behavior of graphene can be described by conserved quantities such as momentum density and energy density. Due to the electronic nature of this mode, the estimated propagation velocity can be 10^3 times larger than that for previously studied phonon mechanisms (phonon second sound). The energy mode is uncharged at charge neutrality, but becomes coupled to charge dynamics upon doping. As a result, this coupling can be used for all-electric excitation and detection of energy transport.

Chapter 2

Hot Carrier Transport and Photoresponse in Graphene

Graphene possesses a unique combination of optoelectronic characteristics [5]over a wide range of frequencies [6, 52]. Its high mobility enables high-speed photodetection [7,8]. Combined with optical transparency and gate-tunable carrier density (field effect), this makes graphene an attractive material for photonic and optoelectronic applications [53, 54].

A defining characteristic of a material's optoelectronic response is photocurrent generated (eg. in p-n type junctions shown in Fig. 2.1b). While strong photocurrent response was observed in graphene [7,8,19,37,55,56], the generation mechanism was debated. As we will see below, hot carriers and photocurrent response in graphene are intimately related; graphene's photoresponse will serve as an ideal setting in which to illustrate how hot carriers have a strong impact on graphene's properties. In this chapter[1], we describe how photocurrent is generated in graphene. In Sec. 2.1 we lay out our theoretical prediction of a hot carrier dominated photoresponse in graphene and in Sec. 2.2 we describe a theory-

[1]The first part of this chapter is reproduced in part with permission from JCW Song, MS Rudner, CM Marcus, LS Levitov, Hot Carrier Transport and Photocurrent Response in Graphene, *NanoLetters* **11** 4688-4692 Copyright 2011 American Chemical Society.

experiment collaboration that confirmed this [2] .

One of the reasons why graphene photoresponse received significant attention was the possibility of harnessing Auger-type processes in which electron-electron interactions allow multiple-excitation generation by a single photon [29, 57, 58]. Multiple excitation generation, known as carrier multiplication for semiconductors, is of fundamental importance for optoelectronics, since many optoelectronic devices can achieve much higher efficiencies if operated in the multiple excitation generation regime[3].

Recent theory suggests that multiple excitation generation can be readily achieved in graphene [29,57,58]. Multiple excitation production can result from Auger-type processes induced by photoexcited carriers [25] (Figure 2.1a shows an intraband Impact Excitation process). Indeed, as we discuss in detail in chapter 4, multiple excitation generation in doped graphene is dominated by electron-electron scattering events within a single band, allowing the primary high energy electron-hole pair created by absorbing a photon to heat up the ambient carrier distribution. This is characterized by fast carrier-carrier scattering, which dominates over electron-phonon scattering in the photoexcitation cascade yielding a large fraction α of the initial photon energy ϵ_0 that is captured by the ambient electronic system (see also Chapter 4) [4].

[2]The figures shown in Sec. 2.2 have been reproduced from NM Gabor, JCW Song, *et. al*, Hot Carrier-Assisted Intrinsic Photoresponse in Graphene, *Science* **334**, 648 (2011). Reprinted with permission from AAAS

[3]Although multiple excitation generation has been reported in nano-particles such as colloidal PbSe and CdSe quantum dots [59–61], at present the interpretation of these measurements remains controversial [62].

[4]Parenthetical note: We distinguish the multiple excitation generation effect in graphene as 'intrinsic' from the effects in avalanche photodiodes, which operate at relatively high reverse bias, sometimes just below breakdown.

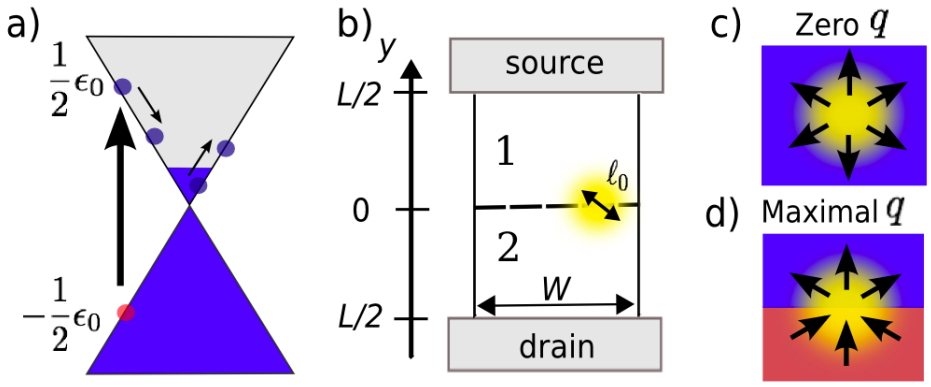

Figure 2.1: (a) Multiple excitation generation in graphene via impact excitation of secondary electron-hole pairs in a single band (see chapter 4 for a microscopic description). (b) Schematic of an optoelectronic device with two separately gated regions 1 and 2, a laser excitation region positioned at the 1-2 interface, and a pair of contacts for collecting photocurrent. (c,d) Photoresponse in homogeneous (c) and inhomogeneous (d) systems. Arrows indicate thermoelectric current due to photogenerated hot carriers. Maximal quantum efficiency is achieved for opposite carrier polarities in regions 1 and 2, indicated by different colors (d).

2.1 Photocurrent Response in Graphene

Here we set out to understand photocurrent response in graphene. We argue that the contribution of multiple excitation generation processes (characterized by the large fraction of energy absorbed, α), which occur locally in the photoexcitation region, is enhanced by energy transport throughout the entire system area. In particular, because of slow electron-lattice relaxation which serves as a bottleneck process for electron cooling [11, 12], photogenerated carriers are thermally decoupled from the crystal lattice over length scales which, even at room temperature, can be as large as several microns. This defines a "hot carrier" regime. Thermoelectric currents, arising in the presence of hot carriers, can lead to a strong photoresponse.

As we shall see, these effects can have a direct impact on the quantum efficiency of photoresponse, namely the number of photogenerated carriers transmitted through the contacts per absorbed photon, $q = N_{\rm el}/N_{\rm ph}$. This quantity is a cumulative characteristic

of the measured photoresponse, which depends on various effects occurring throughout the system, including the multiple excitation generation processes in the excitation region, as well as charge and energy transport from this region to the contacts. These processes are characterized by very different time scales: fast carrier-carrier scattering process is followed by a much longer "charge harvesting" stage dominated by the drift of carriers from the excitation region to the contacts. Here we will focus on the latter stage.

We emphasize that the hot-carrier regime is distinct from the effects of overall heating of both electron and lattice systems. In the latter case, since the electron heat capacity is very small compared to the lattice heat capacity, only a small fraction of the absorbed photon energy, equipartitioned between all degrees of freedom, would remain in the electron subsystem. As discussed in more detail below, this would result in a vanishingly small temperature change, and suppression of the hot-carrier effects.

In contrast, slow electron-lattice relaxation triggers thermal imbalance of the electron and lattice subsystems, amplifying the multiple excitation generation effects. The electron-lattice relaxation slows down for temperatures below the Debye temperature. Under these conditions, the extreme inefficiency of cooling mediated by acoustic phonons allows the carriers to remain hot during their entire lifetime before reaching the contacts. Proliferation of hot carriers dramatically alters the nature of photoresponse.

We exhibit the essential physics of photoresponse by considering a double-gated device comprising two regions with gate-tunable carrier densities (see Figure 2.1b). As we will show, the quantum efficiency q for this system has a simple dependence on the local electrical conductivities $\sigma_{1,2}$ and chemical potentials $\mu_{1,2}$ in the two regions:

$$q = \frac{\alpha\epsilon_0}{(\sigma_1+\sigma_2)^2}\left(\sigma_2\frac{\partial\sigma_1}{\partial\mu_1} - \sigma_1\frac{\partial\sigma_2}{\partial\mu_2}\right), \quad \epsilon_0 = hf, \tag{2.1}$$

where the factor α $(0 < \alpha < 1)$ describes the net fraction of the photon energy ϵ_0 which is

transferred to the electron system through photoexcitation and subsequent decay. Remarkably, under realistic conditions q does not depend on device dimensions and temperature, and can take values as high as $q \gtrsim 2$ (see detailed estimate below).

2.1.1 Slow Electron Cooling and Hot Carriers

We characterize the electron system by an electron temperature, $T_{\rm el}$, which in general is different from the lattice temperature T_0. The electron energy distribution is established via electron-electron scattering which occurs on a sub-picosecond time scale [29, 30, 63]. Since these times are shorter than the electron-phonon timescales, the electron-lattice relaxation can be described by a two-temperature model. Crucially, the processes due to optical phonons, which occur on relatively short times of several picoseconds [64], become *quenched* when the photogenerated carrier energies drop well below the optical phonon energy, $\omega_{\rm D} \approx 200\,{\rm meV}$ [65]. For carriers with lower energies, the dominant cooling process is mediated by acoustic phonons, giving a slow cooling rate which ranging from several picoseconds to a hundred picoseconds [11,12,26–28,66]. For a full discussion of cooling mechanisms see the Introduction and Chapter 3.

Slow cooling results in *thermal decoupling* of the electrons from the crystal lattice and energy transport mediated by hot carriers. As we shall see, the effects due to energy transport dominate over the conventional photovoltaic contribution to photoresponse. The hot-carrier mechanism and the photovoltaic mechanism of photoresponse have very different experimental signatures. In particular, hot carriers manifest themselves in *multiple sign reversals of photoresponse* as carrier concentration is tuned by a monotonic sweep of gate voltage (see Fig.2.2). Multiple sign changes do not occur in the PV effect. Thus the pattern of photocurrent sign changes provides a fingerprint that can be used to experimentally identify the hot-carrier regime.

2.1.2 Hot Carrier Transport

We describe the electric current in the hot-carrier regime through the local current density

$$\mathbf{j} = \sigma\mathbf{E} - e\eta n_{\rm x}(\mathbf{r})\nabla U_g(\mathbf{r}) + \sigma s\nabla T_{\rm el}. \tag{2.2}$$

The first two terms describe the conventional photovoltaic (PV) effect: primary photogenerated carriers are accelerated by the gate-induced electric field $-\nabla U_g$, and create a local photocurrent in the excitation region (here $n_{\rm x}$ is the steady state density of photoexcited carriers and η is the mobility at energy $\epsilon \sim \frac{1}{2}hf$). The redistribution of carriers associated with this local photocurrent sets up an electric field $\mathbf{E} = -\nabla(\phi - \mu/e)$ that drives current outside the excitation region, reaching the contacts. The last term in Eq.(2.2) describes the contribution of energy transferred to electrons via Auger type processes (Fig. 2.1a), which takes the form of a thermoelectric current driven by the electron temperature gradient.

The quantities s and σ in Eq.(2.2) are the Seebeck coefficient and electrical conductivity, which depend on local carrier density and sign. The temperature profile can be found using the energy flux

$$\mathbf{W} = \left(\phi - \frac{\mu}{e}\right)\mathbf{j} - \Pi\mathbf{j} - \kappa\nabla T_{\rm el}, \tag{2.3}$$

where $\Pi = sT$ and κ are the Peltier and the thermal conductivity coefficients. The values of σ, s and κ depend on the microscopic scattering mechanisms. In the practically interesting regime of disorder-dominated scattering, we have

$$s = \frac{\langle(\epsilon-\mu)\tau\rangle}{eT_{\rm el}\langle\tau\rangle}, \quad \kappa = \frac{\sigma}{e^2T_{\rm el}}\left(\frac{\langle\epsilon^2\tau\rangle}{\langle\tau\rangle} - \frac{\langle\epsilon\tau\rangle^2}{\langle\tau\rangle^2}\right), \tag{2.4}$$

where $\langle ...\rangle$ denotes averaging over the energy distribution of carriers, and $\tau(\epsilon)$ is the mean free scattering time for elastic collisions.

We consider a simple model of photocurrent generation in graphene, based on Eqs.(2.2),(2.3), which accounts for the multiscale character of photoresponse: fast carrier

kinetics within a micron-size excitation region set up a pattern of local electron temperature and electric fields that drive current throughout the entire device. As shown in Figure 2.1b, we consider a rectangular graphene sample of width W and length L, with a step in carrier density at the interface between regions 1 and 2. Photocurrent is collected through two contacts placed at $y = \pm L/2$.

In this model, using the continuity relation $\nabla \cdot \mathbf{j} = 0$ and Eq.(2.2), we express the photocurrent as

$$I = \int_0^W \int_{-L/2}^{L/2} \left(s(y) \nabla T_{\rm el} - \sigma^{-1}(y) e\eta n_{\rm x} \nabla U_{\rm g} \right) \frac{dydx}{RW} \tag{2.5}$$

where $R = \frac{1}{W} \int_{-L/2}^{L/2} \sigma^{-1}(y) dy$ is the total resistance, and the contacts are taken to be at equal potentials, $\int_{-L/2}^{L/2} E_y dy = 0$. We focus on the term $s(y)\nabla T_{\rm el}$ in Eq.(2.5), and for the time being ignore the other term. The latter contribution will be analyzed below and shown to be small. Approximating the dependence $s(y)$ by a step that mimics the density profile, we express the hot-carrier (HC) contribution through the average increase of the electron temperature along the 1-2 interface

$$I_{\rm (HC)} = \frac{s_1 - s_2}{R} \Delta T, \quad \Delta T = T^{\rm ave}_{{\rm el},\, y=0} - T_0, \tag{2.6}$$

where s_1 and s_2 are the Seebeck coefficients in regions 1 and 2, T_0 is the lattice temperature, and the superscript 'ave' stands for the value averaged over $0 < x < W$, $y = 0$. The spatial profile of $T_{\rm el}$ must be determined from the heat transport equation

$$\nabla \cdot \mathbf{W} + \gamma C_{\rm el} (T_{\rm el} - T_0) = \alpha \epsilon_0 \dot{N}, \quad \epsilon_0 = hf, \tag{2.7}$$

where γ is the electron-lattice cooling rate, $C_{\rm el}$ is the electron specific heat, and $\dot{N}$ is the photon flux absorbed in the laser spot. Since typical spot sizes $\sim 0.5 - 1\,\mu$m [56] are smaller than other relevant length scales, such as the system size and cooling length (see below), the absorbed photon flux can be approximated as a delta-function source of hot electrons.

2.1.3 Photo-thermoelectric Effect vs Photovoltaic Effect

Next, we analyze the HC photocurrent dependence on gate voltages of regions 1 and 2. Two simple cases are illustrated in Figures. 2.1c,d: the net HC current vanishes for a spatially uniform carrier density, and is maximized for a p-n interface. The full dependence on the chemical potentials $\mu_{1,2}$, illustrated in Figure 2.2a, shows stronger photoresponse for μ_1 and μ_2 of opposite sign. This is in agreement with the 'gate-activated photoresponse' observed in Ref. [67] in the presence of a gate-induced p-n junction.

We model the dependence on the chemical potentials $\mu_{1,2}$, given by the factor $\frac{s_1-s_2}{R}$ in Eq.(2.6), using $R = \frac{L}{2W}\frac{\sigma_1+\sigma_2}{\sigma_1\sigma_2}$ and the Mott formula [68] for the Seebeck coefficient obtained from the non-interacting model (2.4),

$$s(\mu) = -\frac{\pi^2 k_{\rm B}^2 T}{3e}\frac{1}{\sigma}\frac{d\sigma}{d\mu}, \quad \sigma(\mu) = \sigma_{\rm min}\left(1+\frac{\mu^2}{\Delta^2}\right), \tag{2.8}$$

where $k_{\rm B}T \ll \max(\mu, \Delta)$. Here $\sigma(\mu)$ describes a linear dependence of conductivity on carrier concentration away from the Dirac point, with parameters $\sigma_{\rm min}$, the minimum conductivity, and Δ, the width of the neutrality region. The dependence $s(\mu)$, Eq.(2.8), is in good agreement with the measurements of thermopower in graphene [16].

The dependence of photocurrent on μ_1 and μ_2 has a number of interesting features. Because of the dependence on $s(\mu_1) - s(\mu_2)$, the HC current (2.6) vanishes on the diagonal $\mu_1 = \mu_2$ and, in addition, on two hyperbolae $\mu_1\mu_2 = \Delta^2$ that cross the diagonal. We shall refer to the latter as 'anomalous' polarity reversal. As illustrated in Figure 2.2c, this behavior can be traced to the non-monotonic character of the dependence $s(\mu)$. In particular, for any nonzero value of μ_1 excluding extrema of $s(\mu)$, the dependence on μ_2 features two polarity changes. This is illustrated by the slice $\mu_1 = \mu_*$ in Figure 2.2c. At the nodes $\mu_2 = \mu', \mu''$, the Seebeck coefficient satisfies $s(\mu_*) = s(\mu') = s(\mu'')$, as shown by the horizontal dashed line in Figure 2.2c. Hence the photocurrent has opposite polarity

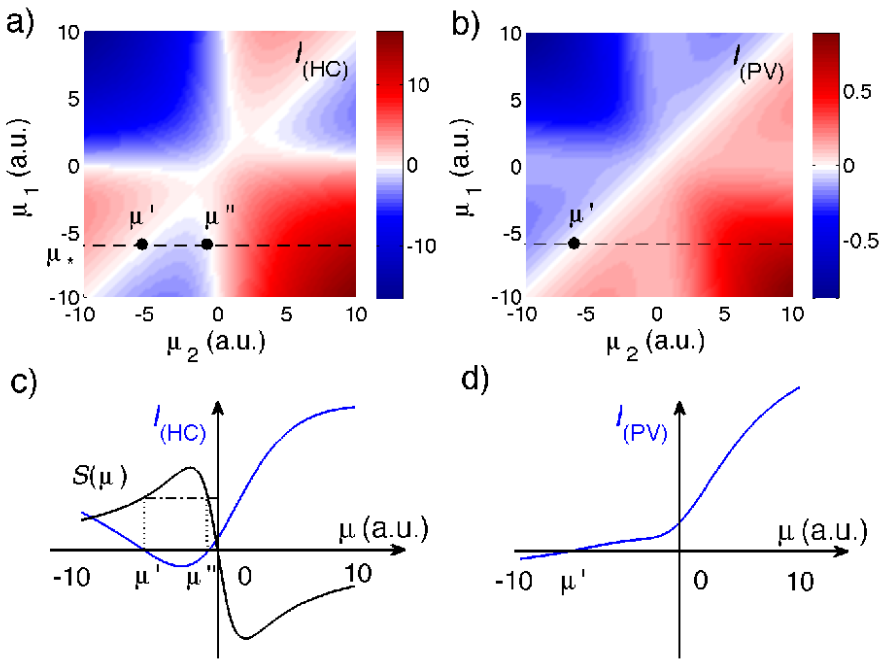

Figure 2.2: Photocurrent map as a function of chemical potentials μ_1 and μ_2 for the device shown in Figure 2.1b. Separately shown are the HC contribution (a) and the PV contribution (b), described by Eq.(2.6) and Eq.(2.10), respectively. Note multiple polarity changes for the HC contribution as opposed to a single polarity change for the PV contribution. The scales of $I_{(\mathrm{HC})}$ and $I_{(\mathrm{PV})}$ have been calibrated to agree with the ratio calculated in Eq.(2.16) with $L = 6\,\mu$m. Current slices along the dotted lines $\mu_1 = \mu_*$ shown for the HC contribution (c) and the PV contribution (d). The Seebeck coefficient $s(\mu)$ is shown in (c).

inside and outside the interval $\mu' < \mu_2 < \mu''$. Using a numerical evaluation of the integral in Eq.(2.6), we checked that multiple polarity reversal, 'proper' and 'anomalous', as well as other qualitative features are insensitive to the photoexcitation spot size, surviving even for spatially uniform photoexcitation.

A model similar to Eq.(2.6) was used in Ref. [56] to describe a laser-induced photothermoelectric effect observed in a heterogeneous system, a monolayer-bilayer interface. The measured photocurrent sign was consistent with Eq.(2.6) but not with the PV effect (see below).

The multiple polarity reversal in the dependence on μ_1 and μ_2 is unique for the HC mechanism. It is instructive to make a comparison with the photocurrent response in the conventional photovoltaic (PV) regime where the primary photogenerated pair is the main contributor to photoresponse. This contribution is described by the second term of Eq.(2.5), giving

$$I_{(\mathrm{PV})} = -\frac{1}{RW}\iint \sigma^{-1}(\mathbf{r}) e\eta n_{\mathrm{x}}(\mathbf{r}) \nabla U_g(\mathbf{r})\, dxdy. \tag{2.9}$$

The integration simplifies when the size of the photoexcitation spot is larger than the depletion length, $l_0 \gg w_d$. Setting $eU_g(\mathbf{r}) = \mu(\mathbf{r})$, using the model dependence $\sigma(\mu)$ from Eq.(2.8), and replacing the integration over y by integration over μ, gives

$$I_{(\mathrm{PV})} = \frac{\eta\Delta}{\sigma_{\min}R}\left(\tan^{-1}\frac{\mu_1}{\Delta} - \tan^{-1}\frac{\mu_2}{\Delta}\right) n_{\mathrm{x}}^{\mathrm{ave}}(y=0). \tag{2.10}$$

Result (2.10) also shows photoresponse maximized in the presence of a p-n junction, i.e. for μ_1 and μ_2 of opposite sign. Thus the 'gate-activated photoresponse' of Ref. [67] cannot be used to distinguish the HC and PV contributions to photocurrent.

The difference between the two contributions to photocurrent is most striking when μ_1 and μ_2 have equal signs [see Figure 2.2a,b]. Since the polarity of the PV current is determined solely by the sign of field gradient ∇U_g, there is only one sign reversal occurring

at $\mu_1 = \mu_2$. In contrast, the nonmonotonic character of $s(\mu)$ produces multiple polarity reversal of the HC contribution. Thus the polarity of photocurrent as a function of gate potentials offers a direct way to differentiate between the two mechanisms.

To estimate the magnitude of the photocurrent, Eq.(2.6), we need to obtain the steady state profile of $T_{\rm el}$ from Eq.(2.7). Since $\mathbf{j}$ has zero divergence, we can write

$$\nabla \cdot \mathbf{W} = -\sigma^{-1}|\mathbf{j}|^2 - \mathbf{j} \cdot \nabla\Pi - \nabla \cdot (\kappa \nabla T). \tag{2.11}$$

The first term, which is quadratic in $\mathbf{j}$, can be ignored. The second term describes the Peltier cooling effect due to photocurrent passing through the 1-2 interface. Incorporating it in Eq.(2.7) gives

$$-\nabla \cdot (\kappa \nabla T_{\rm el}) + \gamma C_{\rm el}(T_{\rm el} - T_0) = \alpha\epsilon_0 \dot{N} + \mathbf{j} \cdot \nabla\Pi. \tag{2.12}$$

Since the spatial extent of the Peltier term $\mathbf{j} \cdot \nabla\Pi$ is of order of the the p-n junction width, which is $\lesssim 0.1\,\mu$m in the state-of-the-art devices, it can be well approximated as a delta function source localized at the 1-2 interface.

Eq.(2.12) can be conveniently analyzed using quantities averaged over the device width $0 < x < W$, $T_{\rm el}^{\rm ave}(y) = \frac{1}{W}\int_0^W T_{\rm el}(x,y)dx$, $\dot{N}^{\rm ave}(y) = \frac{1}{W}\int_0^W \dot{N}(x,y)dx$, and transforming Eq.(2.12) to a one-dimensional equation. For simplicity, we will consider the case when the laser spot is positioned on the 1-2 interface. Treating it as a delta function, we solve Eq.(2.12) by piecing together solutions of the homogeneous equation satisfying zero boundary condition at the contacts $y = \pm L/2$, and performing matching of the boundary values at the 1-2 interface.

Estimating the cooling length $\xi = \sqrt{\kappa/\gamma C_{\rm el}}$ we find a large value of several microns [11, 12] which exceeds $L/2$ for typical device dimensions. Eq.(2.12) can thus be solved by approximating $\gamma \approx 0$, yielding temperature profile $\delta T_{\rm el}^{\rm ave}(y) = (1 - 2|y|/L)\Delta T$ with the

average temperature increase at the 1-2 interface

$$\Delta T = \frac{\alpha\epsilon_0 l_0 \dot{N}^{\rm ave}_{y=0}}{\frac{2}{L}(\kappa_1+\kappa_2)+\frac{T_0}{RW}(s_1-s_2)^2}. \tag{2.13}$$

Since $R \propto L$, ΔT grows linearly with system size, saturating when $L/2$ exceeds the cooling length ξ. For not too high temperatures, $k_{\rm B}T \lesssim \max(\mu_{1,2}, \Delta)$, the second term in the denominator is smaller than the first term and can thus be ignored. Combining this result with Eq.(2.6), we evaluate the quantum efficiency as $q = I_{\rm (HC)}/(el_0 W \dot{N}^{\rm ave}_{y=0})$. Using the Wiedemann-Franz (WF) relation $e^2\kappa = \frac{\pi^2}{3}k_{\rm B}^2 T\sigma$, and the Mott formula, Eq.(2.8), we arrive at Eq.(2.1). The result (2.1) describes the realistic situation of large cooling length, $\xi \gtrsim L$.

For a general system size L, the solution can be obtained as $\delta T(y<0) = A_1 \sinh((y+\frac{1}{2}L)/\xi_1)$, $\delta T(y>0) = A_2 \sinh((\frac{1}{2}L-y)/\xi_2)$. After matching the boundary values and derivatives at $y=0$, we obtain

$$\Delta T = \frac{\alpha\epsilon_0 l_0 \dot{N}^{\rm ave}_{y=0}}{\frac{\kappa_1}{\xi_1}\coth\frac{L}{2\xi_1}+\frac{\kappa_2}{\xi_2}\coth\frac{L}{2\xi_2}+\frac{T_0}{RW}(s_1-s_2)^2}. \tag{2.14}$$

This result agrees with Eq.(2.13) for small system size $L \ll \xi_{1,2}$. At large L, it describes saturation to the value $\Delta T = \alpha\epsilon_0 l_0 \dot{N}^{\rm ave}_{y=0}/(\kappa_1/\xi_1+\kappa_2/\xi_2)$.

For a prelimanary estimate of the numerical value of q, we can use for example the factor α calculated in Ref. [57] which predicts $\alpha\epsilon_0 = Mk_{\rm B}T_0$ with $M = 4.3$, $T_0 = 300\,{\rm K}$, thus giving $\alpha \approx 0.07$. [5] Taking the chemical potentials at the minimum and maximum of $s(\mu)$ described by the model (2.8), $\mu_{1,2} = \pm\Delta$, we obtain $q = Mk_{\rm B}T_0/2\Delta$. Taking a typical value for the neutrality region width for graphene on BN substrate, $\Delta \lesssim 300\,{\rm K}$, we find $q \gtrsim 2$. Thus, high quantum efficiencies are feasible for realistic system sizes of up to

[5] The factor, α, depends intricately on microscopic processes; the exact value of α is the subject of ongoing research in the field and can depend on the energy of the photoexcited carriers, and doping in graphene. Indeed, the microscopic processes leading to α is treated in detail in Chapter 4 where large $\alpha \gtrsim 0.7$ are found both theoretically and inferred from an experiment for a *doped* graphene system [30]. The value $\alpha \approx 0.07$ used here is obtained from Ref. [57] for an *undoped* system which represents a lower bound (conservative estimate) for α.

$5 - 10\,\mu$m. How much energy is captured by the electronic system under photoexcitation depends intricately on the microscopic processes involved, in particular, on the competition between fast carrier-carrier and carrier-optical phonon scattering and affects the factor α. This is discussed in Chapter 4.

We now proceed to estimate the relative strength of the HC and PV contributions to photoresponse. Using the WF relation and the Mott formula, Eq.(2.8), we find

$$\frac{I_{(\mathrm{HC})}}{I_{(\mathrm{PV})}} \approx \frac{\sigma_{\min} s l_0 L \alpha \epsilon_0 \dot{N}^{\mathrm{ave}}}{2\Delta\kappa\eta n_{\mathrm{x}}^{\mathrm{ave}}} = \frac{e\alpha\epsilon_0 l_0 L}{4\eta\tau_0\Delta} \times \frac{\sigma_{\min}}{\sigma^2}\frac{d\sigma}{d\mu}, \tag{2.15}$$

where we estimated the photoexcited carrier density as $n_{\mathrm{x}} = 2\tau_0\dot{N}$ with τ_0 the carrier lifetime. Near the Dirac point, estimating $\frac{\sigma_{\min}}{\sigma^2}\frac{d\sigma}{d\mu} \sim 1/\Delta$, we obtain

$$\frac{I_{(\mathrm{HC})}}{I_{(\mathrm{PV})}} \approx \frac{e\alpha\epsilon_0 l_0 L}{4\eta\tau_0\Delta^2} \approx 2.6L\,[\mu\mathrm{m}] \sim 15 - 25 \tag{2.16}$$

for $L \sim 5 - 10\,\mu$m, where we used parameter values: $\eta = 10^4\,\mathrm{cm}^2V^{-1}s^{-1}$, $\tau_0 = 1\,\mathrm{ps}$ [69], the neutrality region width $\Delta \approx 100\,\mathrm{meV}$ estimated for graphene on SiO substrate [70], $\alpha \approx 0.07$ [57], the photon energy $\epsilon_0 \sim 1.5\,\mathrm{eV}$, and the laser spot size $l_0 \approx 1\,\mu$m [56].

We therefore conclude that, due to very inefficient electron-lattice cooling and an efficient multiple excitation generation process, an abundance of hot carriers leads to a dominant HC contribution to the photoresponse. Furthermore, because the ratio $I_{(\mathrm{HC})}/I_{(\mathrm{PV})}$ scales inversely with the square of the neutrality region width Δ^2, we expect hot-carrier-related phenomena such as high quantum efficiency to become more pronounced for high mobility samples, e.g. graphene on a boron nitride substrate.

To better understand the relation between the result (2.16) and the way energy is partitioned during the photoexcited cascade, it is instructive to consider the situation when energy is equally partitioned between all degrees of freedom, electrons and lattice (which would be the case for very fast electron-lattice relaxation). Crucially, the large difference between the electron and phonon specific heat values makes the lattice act as a nearly ideal

heat sink. In this case, the value $\alpha \approx 0.07$ used above, describing the fraction of photon energy remaining in the electron subsystem, would be replaced by the heat capacity ratio $C_{\rm el}/C_{\rm ph}$, which is very small. In the idealized case of a sharp Dirac point and undoped system, we have $C_{\rm el}/C_{\rm ph} = v_{ph}^2/v_{\rm F}^2 \approx 10^{-4}$. Although $C_{\rm el}$ is somewhat enhanced at finite doping, under realistic conditions it remains quite small. We estimate $C_{\rm el}/C_{\rm ph} \approx (v_{ph}^2/v_{\rm F}^2) \times (\mu/k_BT)$, where $k_BT \ll \mu$. Taking $\mu = 100\,\text{meV}$, for temperatures $10\,\text{K} < T < 300\,\text{K}$ we obtain $C_{\rm el}/C_{\rm ph} \approx 10^{-2} - 10^{-4}$. This would reduce our estimate of the ratio $I_{\rm (HC)}/I_{\rm (PV)}$ by a large factor $10 - 10^3$, strongly suppressing $I_{\rm (HC)}$. Hence, a dominant $I_{\rm (HC)}$ signals that the electronic system and lattice are out of equilibrium.

In summary, hot-carrier transport in the presence of photoexcitation leads to a novel type of photoresponse dominated by photo-thermoelectric effects. This regime is characterized by clear experimental signatures in the photocurrent response of graphene including

i) Multiple sign reversals of photocurrent over a monotonic sweep of gate voltage.

ii) Temperature at the junction is sensitive to the electronic thermal conductivity, and can be tuned by gate voltage, see Eq. 2.13.

iii) The spatial extent of the photocurrent can as large as several microns and is characterized by the electron-lattice cooling length.

Below we discuss how these signatures manifest in an experiment.

2.2 Observing Hot Carrier-Assisted Photoresponse in Graphene

As discussed above, hot carriers should play a key role in the optoelectronic response of graphene. Yet prior to work described in this chapter, experimental measurements had not clearly determined the photocurrent generation mechanism. Numerous initial studies

suggested that photocurrent generated at graphene-metal contacts [7, 8, 19, 37, 55] results from the photovoltaic (PV) effect, in which a built-in electric field accelerates photogenerated charge carriers to the electronic contacts. More recent studies attributed photocurrent in monolayer-bilayer interfaces [56] and highly controlled chemical potential gradients at gate voltage-controlled p-n interfaces [67] to photothermoelectric (PTE) effects, however their evidence was circumstantial. In particular, while the monolayer-bilayer interface studied in [56] invoked the PTE effect, the specific geometry chosen limited the generality of the conclusions, leaving the intrinsic photoresponse mechanism an open question.

In the following, we briefly discuss a theory-experiment collaboration [21] (triggered by our prediction above) which established the dominance of PTE and the role of hot carriers in the photocurrent response of graphene.

2.2.1 Six-Fold Pattern

The key to detecting the mechanism for photocurrent response in graphene is exploiting the multiple sign changes that occur for the PTE effect in the HC regime in a graphene p-n junction as opposed to the single change for the PV effect as predicted in Fig. 2.2a,b.

In Ref. [21], optoelectronic measurements of dual gate voltage-controlled graphene p-n junction devices in the presence of local laser excitation were performed to determine the intrinsic photoresponse, Fig. 2.3a,b. Dual gate control affords the direct manipulation of the device's chemical potential energy landscape, allowing a probe of photoresponse as a function of junction type; tuning the bottom and top gate voltages, V_{BG} and V_{TG} respectively, allows independent control of carrier density of electrons (n-type carriers) and holes (p-type carriers) in each region [71–73]. By applying voltages of opposite polarity on V_{BG} and V_{TG}, p-n junction at the interface of p- and n-type regions in a single graphene sample can be formed. More information about the experimental setup and methods are available in

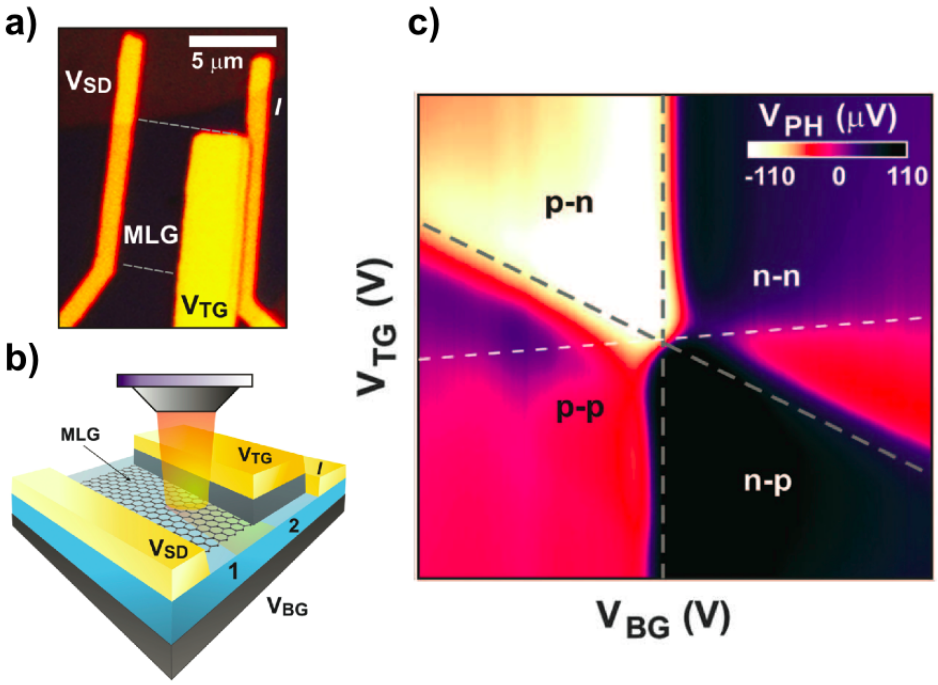

Figure 2.3: a) Optical microscope image of the dual-gated device incorporating boron nitride top-gate dielectric. The current I is measured at fixed bias voltage V_{SD}. Dashed white lines mark the boundaries of graphene. MLG, monolayer graphene; V_{TG}, top-gate voltage. b) Experimental schematic showing laser positioned over p-n junction. Laser position can be moved (see Fig. 2.4c). c) Photovoltage (V_{PH}) versus V_{BG} and V_{TG} at $T = 40\,\mathrm{K}$, measured at the p-n interface. This shows a clear six-fold pattern characteristic of PTE in the Hot Carrier dominated regime, compare with Fig. 2.2. Gray dashed lines are lines of high resistance from transport characteristics. The white dashed line is the line in which carrier density in the two regions are equal. This figure has been adapted from Ref. [21].

Ref. [21]

Focusing a laser ($\lambda = 850\,\mathrm{nm}$) on the junction interface, a striking six-fold photovoltage pattern as a function of gate voltages was found Fig. 2.3. This mimicks the multiple sign changes predicted for the PTE effect in Fig. 2.2a and provides a smoking gun of the PTE and HC regime. As expected from from Eq. 2.6, the multiple sign changes result from the non-monotonic behavior of the Seebeck coefficient, s, as a function of carrier density. Indeed, this same pattern was found for a wide range of temperatures ranging from room temperature down to 10 K and in the linear optical power regime [21].

The tilt in the data shown in Fig. 2.3c as compared with the theoretical prediction

in Fig. 2.2a is due to the geometry of the back gates that control the density in the two regions of Fig. 2.3b: V_{BG} affects the carrier density in the full sample, but V_{TG} only affects the carrier density in a local region so that the carrier density in the two regions n_T and n_B are

$$n_B = C_{\mathrm{BG}} V_{\mathrm{BG}}, \quad n_T = C_{\mathrm{TG}} V_{\mathrm{TG}} + C_{\mathrm{BG}} V_{BG} \tag{2.17}$$

where C_{BG} (C_{TG}) are the bottom (top) gate capacitance to graphene device, which in this device had a ratio $C_{\mathrm{BG}}/C_{\mathrm{TG}} \sim 0.05$. As a result of this gate geometry, the photovoltage plot is skewed in the V_{BG} vs V_{TG} plots shown.

Since this first measurement, multiple groups have observed similar six-fold patterns in dual gate controlled graphene p-n jucntions [27, 74].

2.2.2 Extracting the Seebeck Coefficient

To highlight the thermoelectric origin of the photocurrent in graphene, we can use Fourier analysis to extract the gate-voltage dependence of the photovoltage (or photocurrent) from our V_{BG} vs. V_{TG} plots. This is motivated by the following observation of scalar functions of 2 variables: Consider $f = f(\mathbf{a} \cdot \mathbf{x})$ where $\mathbf{a} = (a_1, a_2)$ is a constant and $\mathbf{x} = (x_1, x_2)$ encode the 2 variables. The Fourier transform of $\tilde{f} = \int d^2\mathbf{x} f(\mathbf{a} \cdot \mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}}$

$$\tilde{f} = \int dx^{\parallel} dx^{\perp} f(|a|x^{\parallel}) e^{-i(q^{\parallel}x^{\parallel} + q^{\perp}x^{\perp})} = 2\pi\delta(q^{\perp}x^{\perp})\tilde{f}(x^{\parallel}/|a|) \tag{2.18}$$

where $x^{\parallel}, q^{\parallel}$ and $x^{\perp}, q^{\perp}$ are components of $\mathbf{x}, \mathbf{q}$ parallel and perpendicular to $\mathbf{a}$ respectively. This means that the 2-d Fourier transform (with respect to $\mathbf{x}$) should exhibit a streak in $\mathbf{q}$ along the direction parallel to $\mathbf{a}$.

Similarly, since the photovoltage arises from the difference of seebeck coefficients in two regions, $V_{\mathrm{PH}} = [s(n_T) - s(n_B)]\Delta T$, the fourier transform of $\tilde{V}_{\mathrm{PH}} = \int d^2\mathbf{v} V_{\mathrm{PH}} e^{-i\mathbf{q}\cdot\mathbf{v}}$

should exhibit two streaks,

$$\tilde{V}_{\mathrm{PH}} \propto \tilde{s}(|q_T^{\parallel}|)\delta(q_T^{\perp} \cdot v_T^{\perp}) + \tilde{s}(|q_B^{\parallel}|)\delta(q_B^{\perp} \cdot v_B^{\perp}) \tag{2.19}$$

where $\mathbf{v} = (V_{\mathrm{BG}}, V_{\mathrm{TG}})$, $\tilde{s}$ is the 1-d fourier transform of the seebeck coefficient, $q_T^{\parallel(\perp)}$ and $v_T^{\parallel(\perp)}$ are $\mathbf{q}$ and $\mathbf{v}$ components parallel (perpendicular) to n_T and $q_B^{\parallel(\perp)}$ and $v_B^{\parallel(\perp)}$ are $\mathbf{q}$ and $\mathbf{v}$ components parallel (perpendicular) to n_B, see Eq. 2.17.

Taking a Fourier transform of V_{PH} we find two streaks as shown in Fig. 2.4a along directions parallel to N_{B} and N_{T} as expected from Eq. 2.19. Indeed, these streaks are perpendicular to the gray dashed lines in Fig. 2.3b which trace out the lines of minimum resistance; the angle come from the capacitive coupling in Eq. 2.17.

Since the amplitudes of each of the steaks is proportional to $\tilde{s}$, we can mask each of the streaks individually and perform an inverse Fourier transform to obtain the gate voltage dependence of the s, up to a multiplicative pre-factor. Performing this analysis in *matlab* [6], we obtain the seebeck coefficient gate voltage dependence under each of the gates in Fig. 2.4b which displays the non-monotonic gate voltage dependence (lower panel) for thermoelectric power in graphene, Eq. 2.8. Indeed, taking the sum of the extracted seebeck coefficients Fig. 2.4b (top panel) in the two regions produces the six-fold pattern observed Fig. 2.4a lower inset.

2.2.3 Electron-Lattice Cooling Lengths

The hot electron temperature is governed by energy transport given in Eq. 2.7. Since the photocurrent is sensitive to the temperature at the junction interface, the spatial profile of

[6]The procedure for masking was as follows: We mask the horizontal streak using a one-dimensional Gaussian (with variance var $= 1/(0.07V)$) and then perform an inverse Fourier transform, shifting the origin back to the sides (resulting in $V_{\mathrm{PH}}[B]$ of Fig. 2.4b). To isolate the diagonal streak, we next take the masked horizontal streak and subtract it from the total Fourier transform. This acts to remove the overlapping elements near the origin. We take a diagonal mask using the same procedure with a slightly larger variance (var $= 1/(0.16V)$), resulting in the $V_{\mathrm{PH}}[T]$ shown in Fig. 2.4b.

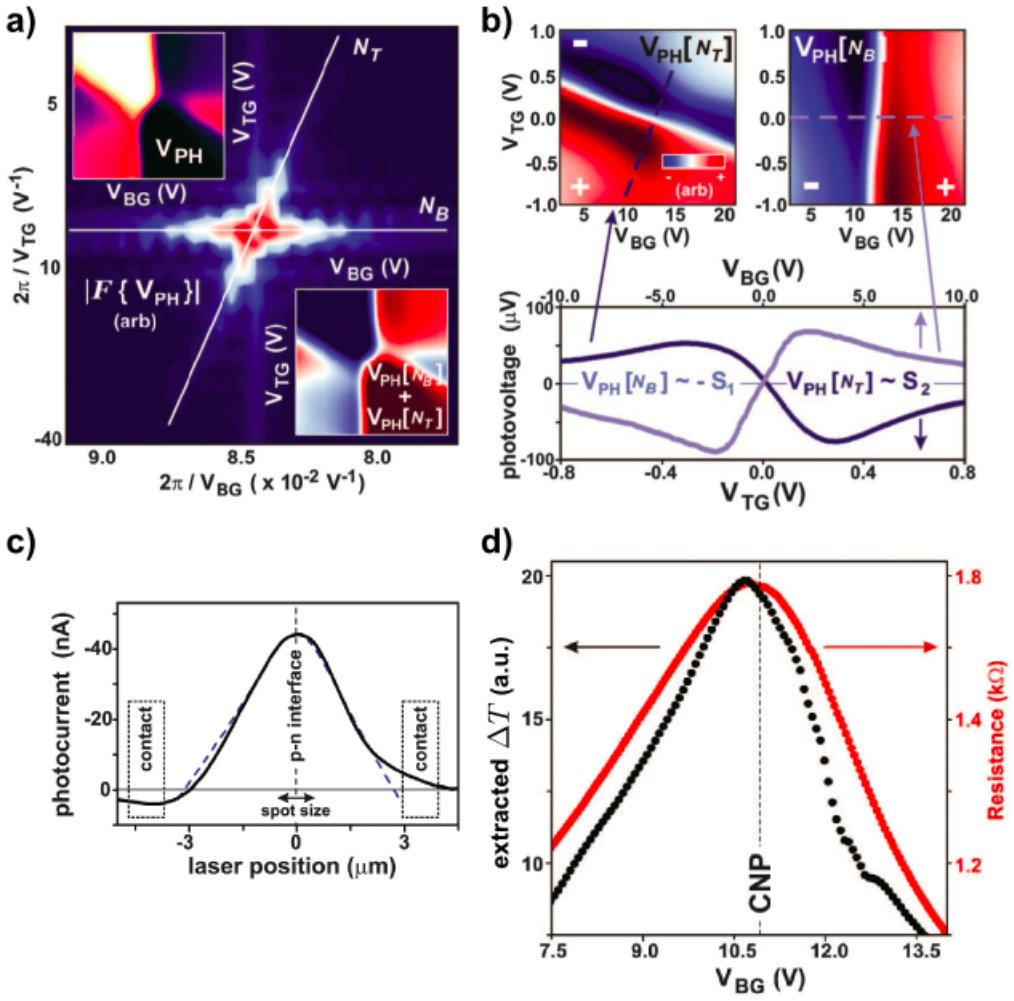

Figure 2.4: a) Fourier analysis of the photovoltage response.Fourier transform $|F\{V_{\mathrm{PH}}\}|$ of photovoltage versus V_{BG} and V_{TG} from Fig. 2.3C. N_{B} and N_{T} label two components of the Fourier transform masked to extract the bottom- and top-gated photovoltage components. Upper inset: Photovoltage map from Fig. 2.3c. Lower inset: Sum of the individual Fourier components of the photovoltage. b) Bottom- and top-gated photovoltage components $V_{\mathrm{PH}}[N_{\mathrm{T}}]$ and $V_{\mathrm{PH}}[N_{\mathrm{B}}]$ as a function of V_{BG} and V_{TG} calculated by inverse Fourier transforming along the masked directions N_{T} and N_{B}. Lower panel: Photovoltage line traces from $V_{\mathrm{PH}}[N_{\mathrm{B}}]$ and $V_{\mathrm{PH}}[N_{\mathrm{T}}]$. The V_{BG} axis has been shifted by the CNP voltage for $V_{\mathrm{PH}}[N_{\mathrm{B}}]$. c) Photocurrent as a function of laser position (see Fig. 2.3 for a schematic). The triangular shape of the photocurrent profile indicates long cooling lengths. (d) Extracted temperature at the p-n interface ΔT, and resistance as a function of V_{BG} measured along the white dashed line in 2.3C. The extracted temperature was obtained by taking the ratio of the photovoltage to the thermovoltage that was measured in the same device. The procedure is outlined in Ref. [21]. This figure has been adapted from Ref. [21].

the hot electron temperature can be extracted from the spatial profile of the photocurrent as the laser source is spatially scanned about the device. This is given by the solution to the heat equation

$$-\nabla\kappa\nabla T_{\rm el} - \gamma C_{\rm el}(T_{\rm el} - T_0) = J_{\rm laser} \tag{2.20}$$

where $J_{\rm laser}$ is the rate of heat absorbed by the electronic system from the laser source. We can treat this as an effectively one dimensional problem and average over the device width. In the finite system length L with zero boundary conditions at the contacts $y = \pm L/2$, the solutions of Eq. 2.20 can be found using the Green's function method. The finite system size Green's function for this is given by

$$\begin{aligned} G(y > y') &= \frac{-\xi}{\kappa}\frac{\sinh([y - L/2]/\xi)\sinh([y' + L/2]/\xi)}{\sinh(L/\xi)}, \\ G(y < y') &= \frac{-\xi}{\kappa}\frac{\sinh([y + L/2]/\xi)\sinh([y' - L/2]/\xi)}{\sinh(L/\xi)}, \end{aligned} \tag{2.21}$$

where y is the distance from the position of the laser, L is the full length of the device, and $\xi = \sqrt{\kappa/\gamma C_{\rm el}}$ is the cooling length. We can obtain the temperature profile for a given cooling length, ξ, by taking the laser spot as a Gaussian and convoluting with the Green's function. As shown in Fig. 2.4c, various fits of the photocurrent temperature profile obtained and can discern that $\xi = 2, 3, \infty$ all fit the measured curve very well.

When cooling lengths approach $L/2$ (or larger than $L/2$), the Green's functions approach a simple piece-wise linear expression $G(y, y' = 0) = (1 - 2|y|/L)\Delta T$, where the constant ΔT is determined by can be determined by Eq. 2.13. Indeed, the triangular shape of the photocurrent spatial profile approximates the piecewise linear function for temperature profile discussed above.

From the cooling length $\xi > 2\,\mu$m, we estimated a lower bound of the hot carrier cooling time $\tau \sim 100\,$ps at $T = 40\,$K. This cooling time is consistent with very slow cooling dynamics recently observed in graphene [66]. The dimensions of the device $L \approx 6\,\mu$m allow

us to establish only a lower bound for the cooling length.

2.2.4 Gate-Tunable ΔT

Graphene's intrinsic photoresponse also exhibits charge density dependence that confirms hot carrier-assisted transport of thermal energy. For excitation fixed at the p-n interface, the steady state value of the electronic temperature increase ΔT at the laser spot is determined by Eq. 2.13 (in the limit of long cooling length). Suppressing the Pelteir contribution (second term in the denominator as small)

$$\Delta T = \frac{J_{\text{laser}}}{K_1 + K_2} \tag{2.22}$$

where dQ/dT is the rate of heat entering the electronic system. Importantly $\kappa_{1,2}$ is the *electronic* thermal conductance which obeys the Wiedemann-Franz relation, $e^2K = \pi^2 k_B^2 T/3 \times (1/R)$, where R is the resistance of the device. As a result, in the hot carrier regime $\Delta T \propto 1/K \propto R$; gate dependent ΔT provides evidence of ΔT of an electronic origin, conversely gate independent ΔT would show that both the electron and lattice systems are being heated up.

We isolated ΔT by analyzing V_{PH} being careful to eliminate any spurious gate dependencies that arise from other gate dependent quantities such as the seebeck coefficient. For a detailed description of the procedure employed, see Ref. [21]. ΔT obtained in this fashion has density dependence shown in Fig. 2.4c (black line) that mimics the density dependence of the device resistance (in red). The strong density dependence peaking at the charge neutrality points clearly demonstrates the electronic origin of ΔT.

2.3 Summary

The clear experimental confirmation (described above in section 2.2) of our prediction that hot carrier effects dominate the photoresponse of graphene (described in section 2.1) illustrates the important role that hot carriers play in graphene's energy flow and optoelectronic response. Indeed, many other groups have also seen the six-fold pattern [27,74], and it is now used as a standard technique for differentiating the hot carrier assisted Photothermoelectric effect in graphene p-n junctions.

Transport of hot electronic carriers results in a novel nonlocal transport regime that may enable increased power conversion efficiency in energy-harvesting devices. At present, graphene is considered to be an excellent candidate for energy-harvesting optoelectronics, in part because of the presence of high responsivity photodetection with high internal quantum efficiency [7]. Hot carrier-assisted photoresponse is predicted to improve the power conversion efficiency of energy-harvesting devices beyond standard limits [34]. These findings together with new design principles for next-generation solar thermoelectric devices, make graphene-based systems viable for energy-harvesting applications.

Chapter 3

Supercollisions & Electron-Lattice Cooling in Graphene

A number of interesting and practically useful phenomena arise when slow cooling between electronic and lattice systems results in long-lived hot carriers proliferating over large spatial scales [75]. Energy transport and energy harvesting mediated by hot carriers is utilized in a variety of applications (calorimetry, bolometry, infrared and THz detectors, etc.). Thermal decoupling of electrons from the crystal lattice in most materials takes place at temperatures of order a few kelvin [76]. In contrast, the rates for electron-lattice cooling in graphene are predicted to be very slow in a much wider temperature range [11, 12]. resulting in new optoelectronic and thermoelectric phenomena as well as other hot-carrier effects [20, 21, 66, 77].

The inefficiency of the standard cooling pathways mediated by optical and acoustic phonons [11, 12] stems from the material properties of graphene. The large value of the optical phonon energy, $\omega_0 = 200\,\mathrm{meV}$, quenches the optical phonon scattering channel below a few hundred kelvin[1]; a small Fermi surface and momentum conservation severely constrain

[1]Factors other than disorder often dominate cooling in other systems at $T > T_{\mathrm{BG}}$. For example the relatively small optical phonon energy in GaAs, $\omega_0 \approx 35\,\mathrm{meV}$, makes optical phonons dominate in GaAs-based 2D systems at $T \gtrsim 40\,\mathrm{K}$ [78, 79].

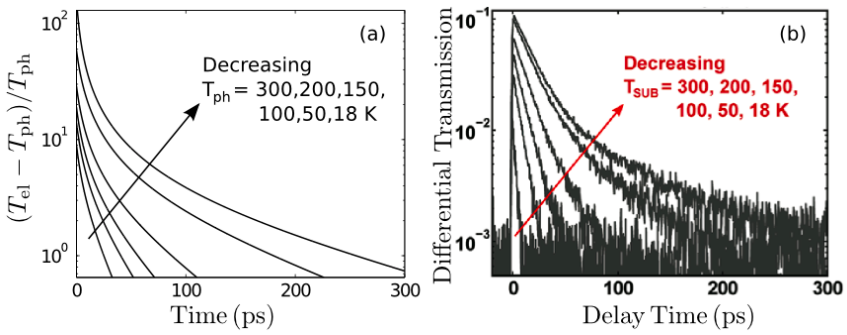

Figure 3.1: (a) Supercollision cooling temperature dynamics in graphene obtained from Eq.(3.1) for the lattice temperature values matching those in panel (b). Parameter values used: doping $\mu = 50\,\text{meV}$ and disorder mean free path $k_F\ell = 20$. (b) Carrier dynamics measured using the pump-probe technique for varying substrate temperatures (reproduced with permission from Fig.2(b) of Ref. [66]).

the phase space for acoustic phonon scattering [11, 12]. Given the momentum conserving character of these processes, which renders them inefficient, cooling in graphene can be particularly sensitive to the effects of disorder. In this chapter[2], we argue that an unconventional, disorder-assisted pathway dominates cooling in a wide range of temperatures, explaining key features of cooling dynamics observed in recent pump-probe measurements [66] (see Fig. 3.1).

The cooling dynamics reported in Ref. [66] features fairly long timescales. The cooling times grow with decreasing temperature, from $\sim 10\,\text{ps}$ at $300\,\text{K}$ to $\sim 200\,\text{ps}$ below $50\,\text{K}$. This is very different from the dependence expected for momentum-conserving scattering by acoustic phonons, where the cooling times are predicted to increase with temperature[3], reaching a nanosecond scale at room temperature for comparable densities [11, 12]. The observed temperature dependence is also clearly distinct from the very steep dependence expected for optical phonons, $\tau \sim e^{\hbar\omega_0/k_\text{B}T}$.

[2]The introduction, first section, and discussion of flexural phonons in this chapter are reproduced from JCW Song, MY Reizer, LS Levitov, Disorder-Assisted Electron-Phonon Scattering and Cooling Pathways in Graphene, *Physical Review Letters* **109**, 106602 Copyright (2012) by the American Physical Society.

[3]see Sec. 3.1.2. This can be understood by noting that the cooling power for single acoustic phonons, J_0 is determined by the size of the Fermi surface, and the heat capacity in the degenerate limit $C_\text{el} \propto T$. As a result, $\gamma \sim \mathcal{J}_0/C_\text{el} \propto 1/T$.

As we show below, a new cooling mechanism (disorder-assisted electron-lattice cooling) yields slow time scales and a temperature dependence that closely match the observations. In addition, as illustrated in Fig.3.1, this mechanism explains subtle features such as the prolonged non-exponential regime of cooling dynamics and the saturation of cooling times at low T manifest in the similarity between the 50 K and 18 K curves [see Eqs.(3.13),(3.14)].

The high impact of disorder on cooling can be understood by noting that the momentum-conserving acoustic phonon processes can only dissipate energy in bits much smaller than k_BT. Indeed, since for such processes the phonon momenta are limited by $2k_F$, the maximal energy transfer cannot exceed $2k_BT_{\rm BG} = 2\hbar sk_F$ per scattering event (here s and k_F are the sound velocity and Fermi momentum). The $T_{\rm BG}$ values are *a few kelvin* for typical carrier densities, i.e. a small fraction of k_BT. In contrast, disorder-assisted scattering (and the supercollision mechanism more generally, discussed below and in Sec 3.3 & 3.4) allows for arbitrarily large phonon recoil momentum values. In this case, the entire thermal distribution of phonons can contribute to scattering, resulting in the energy dissipated per scattering of order k_BT (supercollisions). This provides a dramatic boost to the cooling power. In this chapter, we lay out this new cooling mechanism in graphene.

As described below, modeling disorder by short-range scatterers, we obtain the energy-loss power from this new cooling mechanism as

$$\mathcal{J} = A\left(T_{\rm el}^3 - T_{\rm ph}^3\right), \qquad A = 9.62\frac{g^2\nu^2(\mu)k_{\rm B}^3}{\hbar k_F\ell}, \tag{3.1}$$

where $T_{\rm el}$ ($T_{\rm ph}$) the electron (lattice) temperature, $\nu(\mu)$ is the density of states at the Fermi level per one spin/valley flavor, g is the electron-phonon coupling, and $k_{\rm B}T_{\rm el(ph)} \ll \mu$. The enhancement factor for the disorder-assisted cooling over the momentum conserving

pathways depends on both disorder and temperature:

$$\frac{\mathcal{J}}{\mathcal{J}_0} = \frac{0.77}{k_F\ell}\frac{T_{\rm el}^2 + T_{\rm el}T_{\rm ph} + T_{\rm ph}^2}{T_{\rm BG}^2}, \tag{3.2}$$

[see also Eqs.(3.12),(3.13)]. At room temperature, $T_{\rm el(ph)} \sim 300\,{\rm K}$, and taking $\mu = 100\,{\rm meV}$ ($n \sim 10^{12}\,{\rm cm}^{-2}$) we find $T_{\rm el(ph)}/T_{\rm BG} \approx 50$. For $k_F\ell = 20$, the enhancement factor $\mathcal{J}/\mathcal{J}_0$ can be as large as 100 times.

Given the dominance of the disorder-assisted processes, we predict that cooling in graphene is uniquely sensitive to disorder, $k_{\rm B}T_{\rm el(ph)} \ll \mu$ This sensitivity can account for the wide spread of experimentally measured cooling times [52, 66, 69, 77, 80]. Slow cooling times arise because $\mathcal{J}$ scales linearly with the disorder concentration, via $1/k_F\ell$, and with carrier density, $n \sim \nu^2(\mu)$. The inverse scaling with $k_F\ell$ is consistent with the trend of cooling becoming faster at higher levels of disorder noted in Ref. [52]. The sensitivity to disorder can be used as a knob to engineer cooling rates desirable for specific applications.

The enhancement to phase space may also arise due to processes of other types which we detail below. For example, as we show below flexural phonons can give rise to supercollision processes as well. Indeed, Castro et. al. [81, 82] predicted that scattering by flexural phonons can dominate the resistivity (momentum relaxation) in free standing graphene. In contrast, in this chapter we are concerned with *cooling* which is sensitive to both the scattering rate as well as the exchange in *energy*. For flexural phonons, we find an energy-loss power T dependence resembling that in Eq.(3.1) but with a greatly diminished prefactor. We also discuss how ripples in the graphene sheet (sec. 3.3), and two phonon processes (sec: 3.4) can also contribute to cooling via the supercollision process.

In our discussion of cooling we shall implicitly assume that an effective electronic temperature is established quickly via fast carrier-carrier scattering. This is well justified as carrier-carrier scattering occurs on timescales of tens of femtoseconds [66, 69], far shorter

than the timescales $\gtrsim 1\,\mathrm{ps}$ we are concerned with. Below, we analyze cooling from phonons in the graphene lattice *only*. We note that other phonons (particularly, substrate surface phonons [83]) may also contribute to cooling [84].

3.1 Disorder-Assisted Supercollisions & Electron-Lattice Cooling

We proceed to analyze the disorder-assisted cooling regime, wherein impurity scattering mediates the exchange of momentum and energy between electron and phonon systems. The effect of disorder on electron-phonon scattering can be described by the Hamiltonian

$$\mathcal{H} = \sum_{\mathbf{k},i} \psi^{\dagger}_{\mathbf{k},i} H_0(\mathbf{k}) \psi_{\mathbf{k},i} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} b^{\dagger}_{\mathbf{q}} b_{\mathbf{q}} + \mathcal{H}_{\mathrm{el-ph}} + \mathcal{H}_{\mathrm{dis}}, \tag{3.3}$$

where $H_0 = v_F \sigma \cdot \mathbf{k}$ is the massless Dirac Hamiltonian, identical for $i = 1...N$ spin/valley flavors. The electron-phonon interaction arises from the deformation potential,

$$\mathcal{H}_{\mathrm{el-ph}} = \sum_{\mathbf{q}} g\sqrt{\omega_{\mathbf{q}}}\big(b_{\mathbf{q}} + b^{\dagger}_{-\mathbf{q}}\big) n_{\mathbf{q}}, \quad g = D/\sqrt{2\rho s^2}, \tag{3.4}$$

where D is the deformation potential constant and ρ is the mass density of the graphene sheet.

The transition probability for the emission and absorption of phonons can be described by Fermi's Golden Rule,

$$W_{\mathbf{k}',\mathbf{k}} = \frac{2\pi}{\hbar} \sum_{\mathbf{q}} \left[|M_+|^2 N_{\omega_{\mathbf{q}}} \delta_+ + |M_-|^2 (N_{\omega_{\mathbf{q}}} + 1) \delta_- \right], \tag{3.5}$$

where $\delta_{\pm} = \delta(\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}} \mp \omega_{\mathbf{q}})$, $\mathbf{q}$ is phonon momentum and $N_{\omega_{\mathbf{q}}} = 1/(e^{\beta\omega_{\mathbf{q}}} - 1)$ is the Bose distribution. In the absence of disorder, Eq.(3.4) yields the matrix elements $M^{(0)}_{\pm} = g\sqrt{\omega_{\mathbf{q}}}\delta_{\mathbf{k}'-\mathbf{k}\mp\mathbf{q}}$, where the delta function enforces momentum conservation. In the presence of disorder, possible phonon momenta are unconstrained, taking on any value $|\mathbf{q}| \lesssim q_T = k_B T/s$ [see Fig.3.2 (a)].

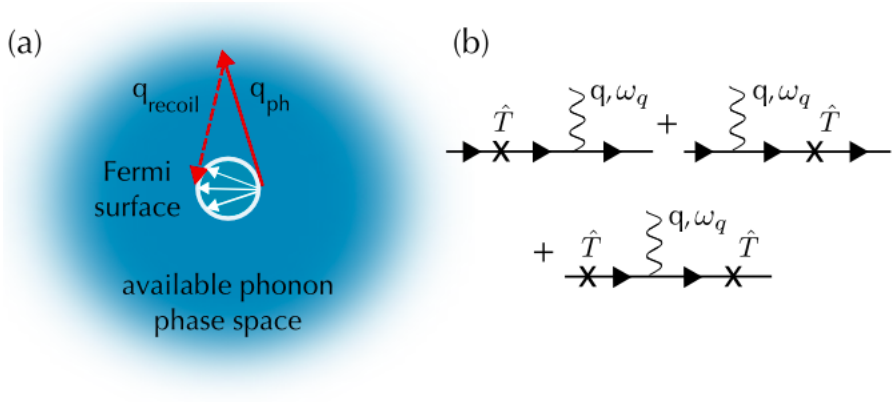

Figure 3.2: (a) Kinematics of supercollisions and normal collisions at $T > T_{\rm BG}$. Phonon momenta are constrained by the Fermi surface for normal collisions (white arrows), and totally unconstrained for supercollisions ($q_{\rm ph}$), with the recoil momentum ($q_{\rm recoil}$) transferred to the lattice via disorder scattering or carried away by second phonon. The energy dissipated in supercollisions is much greater than that dissipated in normal collisions. (b) Feynman diagrams for disorder assisted electron-phonon scattering processes, corresponding to the three terms in Eq.(3.7).

The effect of the disorder potential can be modeled as a sum of impurity potentials and the gauge field originating from strain that couples to electron velocity [85–88],

$$\mathcal{H}_{\rm dis} = \sum_{\mathbf{r},\alpha} \psi^\dagger_\alpha(\mathbf{r}) \left[\sum_i V(\mathbf{r}-\mathbf{r}_i) + v_F \sigma \cdot \mathbf{A}(\mathbf{r}) \right] \psi_\alpha(\mathbf{r}). \tag{3.6}$$

The second type of disorder is peculiar for graphene, where it can arise due to disorder-induced ripples on the graphene sheet [86–90]. The random vector potential $\mathbf{A}(\mathbf{r})$ depends on the microscopic ripple profile. In this section, we will concentrate on the first term; we will consider the contributions from the second term that arises from ripples in the following section.

At low disorder concentration, we can describe disorder-assisted phonon scattering by dressing the electron-phonon vertex with multiple scattering on a single impurity. This gives an expression for the transition matrix elements $M_\pm$ which is exact in the impurity potential:

$$M_\pm = \langle \mathbf{k}' | M^{(0)}_\pm G \hat{T} + \hat{T} G M^{(0)}_\pm + \hat{T} G M^{(0)}_\pm G \hat{T} | \mathbf{k} \rangle, \tag{3.7}$$

where $G(\mathbf{p}) = \frac{1}{\epsilon - H_0(\mathbf{p})}$ is the electron Green's function, $\hat{T}$ is the T-matrix (scattering operator) for a single impurity. The three terms in Eq.(3.7) account for the cases when impurity scattering occurs before or/and after phonon emission (see Fig. 3.2 (b)).

As we shall see, the main contribution to cooling will arise from phonons with momenta of order q_T. Thus we anticipate that the virtual electron states, described by the Green's functions $G(\mathbf{p})$, are characterized by large momenta $|\mathbf{p}| \sim q_T$ which are much greater than $\mathbf{k}$, $\mathbf{k}'$. In this case, for the off-mass-shell virtual states such that $v_F|\mathbf{p}| \gg \mu$, $k_\mathrm{B}T$, we can approximate $G(\mathbf{p}) \approx -\frac{1}{H_0(\mathbf{p})}$. The stiffness of electron dispersion, $v_F \gg s$, along with the estimate $|\mathbf{p}| \sim q_T$, makes it an accurate approximation for all virtual states not too close to the Fermi surface. In this limit, as we now show, drastic simplifications occur because of the particle-hole symmetry $H_0(-\mathbf{p}) = -H_0(\mathbf{p})$.

3.1.1 Short-range Disorder Assisted Supercollisions

We focus on the case of short-range disorder, modeled by a delta function potential $V(\mathbf{r} - \mathbf{r}_j) = u\delta(\mathbf{r} - \mathbf{r}_j)(\hat{1} \pm \sigma_z)/2$, where the plus (minus) sign correspond to impurity positions on the A (B) sites of the carbon lattice. In this case, a nonzero result for the transition matrix element $M_\pm$ is obtained at first order in u. We approximate $\hat{T}_{\mathbf{p}',\mathbf{p}} = \frac{1}{2}u(\hat{1} \pm \sigma_z)e^{i(\mathbf{p}'-\mathbf{p})\mathbf{r}_j} + O(u^2)$ and evaluate the first two terms in Eq.(3.7). This gives the commutator of $H_0^{-1}(\mathbf{q})$ and $\pm\sigma_z$, arising because the virtual electron states in the first and second term have momenta $\mathbf{p} \approx -\mathbf{q}$ and $\mathbf{p} \approx +\mathbf{q}$ (see above)[4]. We obtain

$$M_\pm = \frac{\pm iug\sqrt{\omega_\mathbf{q}}}{\hbar v_F|\mathbf{q}|^2}\langle\mathbf{k}'|(\sigma \times \mathbf{q})_z|\mathbf{k}\rangle e^{i(\mathbf{p}'-\mathbf{p}\pm\mathbf{q})\mathbf{r}_j}, \tag{3.8}$$

[4]The finite commutator, $[\sigma_z, \sigma \cdot \mathbf{q}] = 2i(\hat{\mathbf{z}} \times \mathbf{q}) \cdot \sigma$ arises because of the matrix/sublattice structure of the impurity potential used. As we will see in Sec. 3.3 and Sec. 3.4, this structure is important; for impurities that an identity in sublattice space, M vanishes.

with the phase factor describing the dependence on the impurity position. We evaluate the energy-loss power as

$$\mathcal{J} = \sum_{\mathbf{k},\mathbf{k}',i} W_{\mathbf{k}',\mathbf{k}}(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) f(\epsilon_{\mathbf{k}}) \big[1 - f(\epsilon_{\mathbf{k}'})\big] \tag{3.9}$$

where $f(\epsilon) = 1/(e^{\beta(\epsilon-\mu)} + 1)$ are Fermi functions, $W_{\mathbf{k}',\mathbf{k}}$ is the transition probability, and $\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}$ is the energy exchanged in each scattering event.

In the degenerate limit, $k_B T \ll \mu$, the sum over $\mathbf{k}$ and $\mathbf{k}'$ is conveniently factored into separate integration over energies and angles $\sum_{\mathbf{k},\mathbf{k}'} = [\nu(\mu)]^2 \iint d\epsilon d\epsilon' \iint \frac{d\theta d\theta'}{(2\pi)^2}$. One of the energy integrals is eliminated by the delta functions $\delta(\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}} \pm \omega_{\mathbf{q}})$. The second energy integral is evaluated using the identity $\int_{-\infty}^{\infty} d\epsilon f(\epsilon)(1 - f(\epsilon + \omega_q)) = \omega_{\mathbf{q}} \big[N^{\rm el}_{\omega_{\mathbf{q}}} + 1\big]$, where $N^{\rm el}$ is the Bose distribution function evaluated at the *electron temperature*. With the electron-phonon matrix element given by Eq.(3.8), and using the angle-averaged quantity $\langle |\langle \mathbf{k}'|(\sigma \times \mathbf{q})_z|\mathbf{k}\rangle|^2 \rangle_{\rm ave} = |\mathbf{q}|^2/2$ we obtain an expression[5]

$$\mathcal{J} = \frac{\pi N g^2 u^2}{\hbar^3 v_F^2} [\nu(\mu)]^2 n_0 \sum_{\mathbf{q}} \frac{\omega_{\mathbf{q}}^3}{|\mathbf{q}|^2} \left[N^{\rm el}_{\omega_{\mathbf{q}}} - N_{\omega_{\mathbf{q}}}\right],$$

where n_0 is impurity concentration. Integration yields Eq.(3.1), where we used an expression for the mean free path $k_F \ell = 2\hbar^2 v_F^2/(u^2 n_0)$ [91].

We can make a comparison with the normal (momentum conserving) processes [11, 12], where the cooling power is $\mathcal{J}_0 = B(T_{\rm el} - T_{\rm ph})$, where $B = \pi N \lambda \hbar \nu(\mu) k_F^2 s^2 k_{\rm B}$, and $\lambda = g^2 \nu(\mu)$ is the dimensionless electron-phonon coupling. Linearizing Eq.(3.1) in $\Delta T = T_{\rm el} - T_{\rm ph}$, we find that this contribution dominates over $\mathcal{J}_0$ at temperatures

$$T > T_* = \sqrt{\frac{B}{3A}} = \left(\frac{\pi}{6\zeta(3)} k_F \ell\right)^{1/2} T_{\rm BG}. \tag{3.10}$$

Taking $k_F \ell = 20$ for a rough estimate, we see that the disorder assisted cooling channel dominates for $T \gtrsim 3T_{\rm BG}$. The crossover temperature can be controlled by gate voltage,

[5] We sum over sublattice A(B) of the impurities only after squaring $|M|^2$ [91].

since $T_{\rm BG} \propto \sqrt{n}$. For typical carrier densities n this gives a crossover temperature T_* of a few tens of kelvin.

Interestingly, both at $T < T_*$ and $T > T_*$ the supercollision frequency remains lower than that for normal processes, $\langle W_0 \rangle = \frac{2\pi}{\hbar}\lambda k_{\rm B} T$ [11, 12]. We can define the average collision frequency as

$$\langle W \rangle = \frac{\sum_{\mathbf{k},\mathbf{k}'} W_{\mathbf{k}',\mathbf{k}} f(\epsilon_{\mathbf{k}})\big[1 - f(\epsilon_{\mathbf{k}'})\big]}{\sum_{\mathbf{k}} f(\epsilon_{\mathbf{k}})\big[1 - f(\epsilon_{\mathbf{k}})\big]}. \tag{3.11}$$

Evaluating the integrals as above and setting $T_{\rm el} = T_{\rm ph}$, we find $\langle W \rangle = \frac{4\pi}{\hbar k_F \ell}\lambda k_{\rm B} T \ln \frac{T}{T_{\rm BG}} \ll \langle W_0 \rangle$. The low value for $\langle W \rangle$ means that normal collisions produce the dominant contribution to resistivity even when their contribution to cooling is totally overwhelmed by supercollisions.

3.1.2 Supercollision Temperature Dynamics

Interestingly, the cooling times describing relaxation to equilibrium, $\Delta T_{\rm el}(t) = e^{-(t-t_0)/\tau}\Delta T_{\rm el,0}$, exhibit a nonmonotonic T dependence for $T \sim T_*$. Accounting for both the disorder-assisted and momentum-conserving cooling, the relaxation dynamics can be described as

$$d\mathcal{Q}/dt = -\mathcal{J} - \mathcal{J}_0, \tag{3.12}$$

where $\mathcal{Q}$ is the electron energy density. Taking $\mathcal{Q} = C\Delta T$, with $C = \alpha T$ the heat capacity of the degenerate electronic system ($\alpha = \frac{\pi^2}{3} N\nu(\mu) k_{\rm B}^2$), we find[6]

$$\frac{1}{\tau} = \frac{3A}{\alpha} T + \frac{B}{\alpha T}. \tag{3.13}$$

The cooling time increases with T at $T < T_*$ and decreases at $T > T_*$, reaching maximal value at $T = T_*$. The non-monotonic temperature dependence in τ provides a clear experi-

[6]These cooling times are obtained by linearizing Eq. 3.12 in $\Delta T = T_{\rm el} - T_0$ to obtain $d\Delta T/dt = -(1/\tau)\Delta T$.

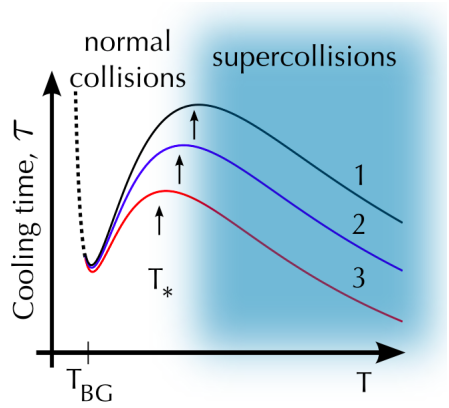

Figure 3.3: Non-monotonic temperature dependence of the cooling time, Eq. 3.13. The up (down) solid line branch is dominated by normal collisions (supercollisions). The three curves illustrate the behavior for different amounts of disorder $A_1 < A_2 < A_3$. The dashed line schematically indicates quenching of cooling at the lowest temperatures ($T < T_{\rm BG}$) [75].

mental signature of the competition between different cooling pathways (shown schematically in Fig. 3.3.)

To describe the cooling dynamics both near and away from equilibrium, we used non-linearized quantities, $\mathcal{Q} = \frac{1}{2}\alpha T_{\rm el}^2$ and Eq.(3.1), with the deformation potential constant $D = 20\,{\rm eV}$ [11, 12], the electron temperature initial value $T_{\rm el,0} = 3 \cdot 10^3\,{\rm K} \sim \omega_0$ and other parameter values cited in Fig.3.1 caption. For the parameters used, $T_* \approx 15\,{\rm K}$. The resulting dynamics, shown in Fig.3.1(a), reproduces the main features seen in the data. We note that the long time behavior is insensitive to the choice of $T_{\rm el,0}$; only the dynamics at short times are affected.

Importantly, there are 2 different regimes for temperature dynamics. First, for short times and $T_{\rm el} \gg T_{\rm ph}$, non-exponential power-law temperature dynamics occurs. At later times when the electronic system has cooled sufficiently so that $T_{\rm el} \approx T_{\rm ph}$, exponential cooling dynamics (described above) sets in. This can be analyzed by non-dimensionalizing Eq. 3.12 and considering the 2 different regimes as detailed below.

The non-exponential behavior seen in the data at short times can be understood by analyzing the regime $T_{\rm el} \gg T_{\rm ph}$. Approximating $\mathcal{J} \approx AT_{\rm el}^3$ and suppressing $\mathcal{J}_0$, we obtain

a $1/(t-t_0)$ dynamics:

$$T_{\rm el}(t) = \frac{T_{\rm el,0}}{1+(A/\alpha)(t-t_0)T_{\rm el,0}}. \tag{3.14}$$

The dynamics at intermediate times, where $T_{\rm el} \gtrsim T_{\rm ph}$, can be found by directly solving Eq.(3.12). We obtain

$$-\frac{2}{\tau}(t-t_0) = F(T_{\rm el}(t)/T_{\rm ph}) - F(T_{\rm el,0}/T_{\rm ph}), \tag{3.15}$$

where $F(x) = 2\sqrt{3}\arctan[(1+2x)/\sqrt{3}] - \ln[(x^3-1)/(x-1)^3]$ (we suppressed the $\mathcal{J}_0$ term which becomes important only for $T \lesssim T_*$ and only at long times). This solution, with τ taken from Eq.(3.13), was used to generate Fig. 3.1 (a), yielding results strikingly similar to the data. This solution is only characterized by two parameters : the initial dimensionless temperature $x_0 = T_{\rm el,0}/T_{\rm ph}$ and the time constant $\tau = 3T_{\rm ph}A/\alpha$ allowing a clear way of comparing with experiments that probe the carrier dynamics of graphene (see section below for discussion of experiments).

In addition, as illustrated in Fig. 3.1, this mechanism explains subtle features such as the prolonged nonexponential regime of cooling dynamics and the saturation of cooling times at low T manifest in the similarity between the 50 and 18 K curves (see Eq. 3.12 and Eq.3.13). We note that the long time behavior is insensitive to the choice of $T_{\rm el,0}$; only the dynamics at short times are affected. Importantly, for supercollision, cooling times increase as temperature decreases as seen in Fig. 3.1.

Besides carrier dynamics, cooling can also be probed by transport measurements through bolometry [22], described by the thermal impedance, $R_{\rm th} = (d\mathcal{P}/d\Delta T)^{-1}$. Here $\mathcal{P}$ is the power pumped into the system (say, via Joule heating), which is balanced by $\mathcal{J}$ in a steady state. Hence, $R_{\rm th}$ temperature dependence can be used as a diagnostic for the processes dominant in cooling. For disorder-assisted cooling, extending Eq.(3.1) to $\mu \lesssim k_{\rm B}T$, we approximate $R_{\rm th}^{-1} = AT^2\big(1+c(k_{\rm B}T/\mu)^2\big)$ for monolayer graphene, with c a constant of

order unity. This temperature dependence is markedly different from $R_{\rm th}$ predicted for momentum conserving channels [92].

3.2 Experimental Observation of Supercollisions

The prediction of the supercollision cooling mechanism above sparked a number of experiments that measured the electron-lattice cooling dynamics and the cooling power of graphene [27, 28]. We briefly comment on their findings.

Graham and co-workers extracted both the steady state electronic temperature and cooling dynamics (in two separate experiments) when irradiated by a continuous wave of light (CW) or pulsed irradiation respectively. In both cases, they used the photothermoelectric effect described in Chapter 2 as a thermometer of the electrons in graphene [27]. Under CW in the steady state, they found that the steady electronic temperature scaled with as a cube root of the amount of power pumped into the system, $T_{\rm el} \propto P^{1/3}$, when $T_{\rm el} \gg T_0$. This is in agreement, with Eq. 3.1 above. Under pulsed illumination, they were able to extract the cooling dynamics of the system using a pump-probe setup [27].

For the cooling dynamics, Ref. [27] found a distinct two regime cooling dynamics. When pumped intensely with pulsed irradiation, the electronic temperature (which was high) initially decayed as a power law similar to that predicted in Eq. 3.14. In particular, as they pumped harder with more photons striking the graphene sheet, the decay rate in this initial regime became faster as anticipated by the rate $AT_{\rm el,0}/\alpha$ scaling linearly with the initial electron temperature in Eq. 3.14. As the electronic temperature decayed, it entered a second regime which could be described by an exponential decay, which is expected from supercollision cooling when $T_{\rm el} \approx T_0$. The CW and pulsed illumination provided different ways in which to measure electronic cooling in graphene. The fact they find agreement with

Eq. 3.1 above over a broad range of electronic and lattice temperatures clearly demonstrates that supercollisions dominates the cooling dynamics of graphene.

Betz and co-workers adopted an alternative approach in which they measured the electronic temperature through noise thermometry [28]. By Joule heating the electronic system, and then subsequently extracting its (steady state) electronic temperature, Betz and co-workers were able to establish that for lattice temperatures larger than T_{BG}, the electronic temperature scaled as $T_{\mathrm{el}} \propto P^{1/3}$ in agreement with Ref. [27] and Eq. 3.1. Further, they extracted A of Eq. 3.1 from their measurements for various dopings (gate voltages) in their sample. They found that it scaled as density or μ^2, consistent with the expectation of Eq. 3.1.

Another interesting prediction of supercollisions is the non-monotonic temperature of cooling time τ as seen in Eq. 3.13 and Fig. 3.3. Below T_* cooling is dominated by the "normal" process in which single acoustic phonons are emitted, and above T_* supercollision dominate; these two regimes exhibit cooling times with different temperature dependences manifesting in a peak of cooling time at T_*. Recently, Ma and co workers [51] attributed a non-monotonic temperature dependence of photocurrent in clean graphene/hexagaonal boron nitride samples[7] to a non-monotonic temperature dependence in cooling lengths, $\xi = \sqrt{\kappa\tau/C_{\mathrm{el}}}$; photocurrent peaked at a temperature $T \sim 50\,\mathrm{K}$. Additionally, the extent of photocurrent profiles (described by cooling length) also peaked at the same intermediate temperature.

These observations [27, 28, 51] confirm that the cooling process in graphene is controlled by supercollisions.

[7] Ref. [51] used a photocurrent technique similar to the one described in Chapter 2.

3.3 Other Disorder-Assisted Supercollisions

While (short-range) disorder-assisted supercollisions described above probably give the largest contribution to electron-lattice cooling at room temperature, the supercollisions mechanism is general and can be applied to a variety of sources of disorder. In this section, we examine other sources of disorder that can give rise to supercollision processess.

3.3.1 Ripple-Assisted Supercollisions

Next, we consider scattering by random strain in the graphene sheet. This type of disorder is described by the vector potential term in Eq.(3.6). Evaluating the electron-phonon matrix element, Eq.(3.7), we find a nonzero contribution at lowest order in $\mathbf{A_q} = \int d^2r e^{-i\mathbf{qr}}\mathbf{A}(\mathbf{r})$. The estimate proceeds in the same way as for short range disorder. The matrix structure $\sigma \cdot \mathbf{A}(\mathbf{r})$ generates the commutator $[\sigma \cdot \mathbf{A_q}, H_0^{-1}(\mathbf{q})] = \frac{-1}{\hbar v_F|\mathbf{q}|} 2i\sigma_z \widehat{\mathbf{q}} \times \mathbf{A_q}$, where $\widehat{\mathbf{q}} = \mathbf{q}/|\mathbf{q}|$, giving

$$M_{\pm}(\mathbf{A}_q) = \langle \mathbf{k}'| \frac{2g\sqrt{\omega_q} i\sigma_z}{\hbar|\mathbf{q}|} \widehat{\mathbf{q}} \times \mathbf{A}_q |\mathbf{k}\rangle. \tag{3.16}$$

Evaluating the transition rate $W_{\mathbf{k}',\mathbf{k}}$ and plugging it into Eq.(3.9), we obtain the cooling power

$$\mathcal{J}_{\text{ripple}} = b \sum_{\mathbf{q}} \omega_{\mathbf{q}} \big(\tilde{N}_{\omega_{\mathbf{q}}} - N_{\omega_{\mathbf{q}}}\big) \langle |\widehat{\mathbf{q}} \times \mathbf{A_q}|^2 \rangle_{\text{dis}}, \tag{3.17}$$

where $b = 4\pi N g^2 s^2 \big[\nu(\mu)\big]^2/\hbar$ and averaging over the ripple ensemble is denoted by $\langle ... \rangle_{\text{dis}}$.

We note that the pair correlator $\chi(\mathbf{q}) = \langle A_{\mathbf{q}} A_{-\mathbf{q}} \rangle_{\text{dis}}$ decays at $|\mathbf{q}| \approx q_0 = 2\pi/R$, where R is the characteristic radius of a ripple [86]. Here we assume that the radius R is much shorter than the phonon wavelength for temperatures of interest (indeed, Ref. [90] estimates $R = 4\,\text{nm}$, whereas at room temperature $\lambda_T = 2\pi/q_T \approx 17\,\text{nm}$)

Since at $q_T \ll q_0$ the integral over $\mathbf{q}$ in the general expression (3.17) is limited solely by the Bose functions $N_{\omega_{\mathbf{q}}}$, the correlator $\chi(\mathbf{q})$ is approximately given by its value at

$|\mathbf{q}| \ll q_0$. This allows us to evaluate the integral over the Bose functions treating $\chi(q)$ as a constant. The correlator $\chi(\mathbf{q})$ was analyzed in Ref. [86]. For the roughness exponent $2H \approx 1$ corresponding to graphene on a substrate [89], Ref. [86] obtains $\chi(|\mathbf{q}| \ll q_0) \approx (\hbar/a)^2 Z^4/R^2$ where Z is the characteristic out of plane displacement in the graphene sheet caused by ripples and a is the interatomic distance. This yields a cooling power

$$\mathcal{J}_{\text{ripple}} = A_{\text{ripple}} \left(T_{\text{el}}^3 - T_{\text{ph}}^3\right), \tag{3.18}$$

where A_{ripple} is

$$A_{\text{ripple}} = \frac{2\zeta(3)k_{\text{B}}^3 b}{2\pi\hbar^2 s^2} \int \frac{d\theta_q}{2\pi} \langle |\tilde{\mathbf{q}} \cdot \mathbf{A}_{|\mathbf{q}| \ll q_0}|^2 \rangle_{\text{dis}} = \frac{\zeta(3) b Z^4 k_B^3}{2\pi s^2 a^2 R^2}, \tag{3.19}$$

where we accounted for a factor of $1/2$ due to angular integral. Comparing this result with the contribution of normal collisions [11, 12], we obtain a crossover temperature

$$T_*^{\text{ripple}} = \eta^{-1/2} T_{\text{BG}}, \quad \eta = \frac{6\zeta(3)}{\pi} \frac{Z^4}{R^2 a^2}. \tag{3.20}$$

Taking the values reported in Ref. [90], $Z = 0.6\,\text{nm}$ and $R = 4\,\text{nm}$, we find $\eta \approx 0.9$, which gives $T_* \approx T_{\text{BG}}$. For a doping level of $\mu = 100\,\text{meV}$ this is $T_*^{\text{ripple}} = 12\,\text{K}$. For weaker disorder, we anticipate T_* to be a few times T_{BG}. We note these are optimistic estimates of ripple assisted supercollision; on cleaner samples of G on hexagonal Boron Nitride, the amount of ripple induced disorder should be far smaller than that estimated above.

3.3.2 Coulomb Disorder-Assisted Supercollisions

Here, we consider the case of Coulomb scatterers. The transition matrix element is described by Eq.(3.7). We can use the Born approximation to write $\hat{T}_{\mathbf{p}',\mathbf{p}} = \frac{2\pi Z e^2}{\kappa|\mathbf{p}'-\mathbf{p}|}$, where Z is the impurity charge and κ is the dielectric constant. We take the unscreened Coulomb potential since the main contribution to cooling come from phonons with momenta $q_T \gg k_F$. Additionally, since Coulomb scatterers are long-ranged, we do not differentiate between

sublattices; any differences arising from having Coulomb impurities on different sites can be absorbed into the scattering from short-range scatterers.

To evaluate the matrix element, we begin with the first two terms of Eq.(3.7). Following the analysis above, we find that the sum of the first two terms of Eq.(3.7) evaluates approximately as the commutator between $G(\mathbf{p})$ and $\hat{T}$. Since $\hat{T}$ for Coulomb scatterers is an identity matrix in graphene's pseudospin space, this commutator *vanishes* and the sum of the first two terms evaluates to zero.

We proceed to consider the last term in Eq.(3.7). Noting that the dominant contributors to cooling arise from phonons with momenta q_T we estimate

$$M_{\pm} = \frac{g\sqrt{\omega_{\mathbf{q}}}}{\kappa^2} \sum_{\mathbf{p}} \langle \mathbf{k}' | \frac{(2\pi Z e^2)^2}{|\mathbf{p}-\mathbf{q}||\mathbf{p}|} H_0^{-1}(\mathbf{p}-\mathbf{q}) H_0^{-1}(\mathbf{p}) | \mathbf{k} \rangle. \tag{3.21}$$

We can evaluate this by rewriting $H_0^{-1}(\mathbf{p})/|\mathbf{p}| = \sigma \cdot \nabla_{\mathbf{p}} \frac{1}{|\mathbf{p}|}$. Performing a 2D Fourier transform (to the conjugate variable $\mathbf{r}$) we obtain

$$M_{\pm} = c \int d^2\mathbf{r} \langle \mathbf{k}' | \frac{(\sigma \cdot \mathbf{r})^2}{|\mathbf{r}|^2} e^{i\mathbf{q}\cdot\mathbf{r}} | \mathbf{k} \rangle = c(2\pi)^2 \delta(\mathbf{q}) \langle \mathbf{k}' | \mathbf{k} \rangle, \tag{3.22}$$

where $c = g\sqrt{\omega_{\mathbf{q}}}(2\pi Z e^2)^2/(\kappa^2 v_F^2 \hbar^2)$. Since the contributions we are interested in are $|\mathbf{q}| \gg k_F$, the matrix element *vanishes*. Hence, we conclude that Coulomb scatterers do not contribute to the supercollision cooling channel.

For convenient comparison, we list the variety of supercollision processes considered in this chapter in Fig. 3.4. As seen in Fig. 3.4, the disorder mediated processes (short-range disorder and ripples) give rise to the main contributions to supercollision cooling.

3.4 Two-Phonon Supercollisions and Electron-Lattice Cooling

It is interesting to consider whether supercollisions can occur in pristine graphene; the emission of two phonons can also give rise to supercollisions. In the following we detail how the cooling powers that arise from such scattering processes.

Supercollision Process	Cooling Power (Linearized)	Crossover Temperature, T_*	Equation
Short-Range Disorder + 1 Acoustic Phonon	$T^2\Delta T$	$\approx$ few T_{BG}	Eq. 3.1
Ripple Disorder + 1 Acoustic Phonon	$T^2\Delta T$	$\approx$ few T_{BG}	Eq. 3.18
Coulomb Disorder + 1 Acoustic Phonon	vanishes	-	Eq. 3.22
Two Acoustic Phonon (vertex)	$T^5\Delta T$	$\approx 20\,T_{\mathrm{BG}}$	Eq. 3.25
Sequential Two Acoustic Phonon	vanishes	-	Eq. 3.26
Two Flexural Phonons	$T^2\Delta T$	$\approx 10\,T_{\mathrm{BG}}$	Eq. 3.28

Figure 3.4: Table of various supercollision processes in monolayer graphene and their linearized cooling power. T_* indicates the crossover temperatures below which the single acoustic phonon emission cooling, $\mathcal{J}_0$, wins over the supercollision process.

3.4.1 Two-Acoustic Phonon Emission Supercollisions

We consider two-phonon scattering which can utilize the supercollision phase space enhancement in much the same way as the disorder-assisted channel (Fig.3.2). This process is described by an anharmonic coupling [93, 94]

$$\mathcal{H}_{\rm el-2ph} = D' \sum_{\mathbf{q_1},\mathbf{q_2}} n_{-\mathbf{q}_1-\mathbf{q}_2} u_{\mathbf{q_1}} u_{\mathbf{q_2}}, \tag{3.23}$$

where $u_{\mathbf{q}} = \sqrt{\omega_{\mathbf{q}}/2\rho s^2}(b_{\mathbf{q}} + b^\dagger_{-\mathbf{q}})$, and D' is an energy of order unity in atomic units. The terms of the form $b^\dagger_{\mathbf{q}_1} b_{\mathbf{q}_2}$ describe Compton scattering of phonons by electrons. These processes have a negligible contribution to the cooling power and can be neglected.

Since $\mathbf{k}$, $\mathbf{k}'$ are on the Fermi surface, and the phonon momenta are large, $q_T \gg k_F$, the two phonons *are nearly counterpropagating*. We can thus set $\mathbf{q_1} \approx -\mathbf{q_2} = \mathbf{q}$. This simplifies the matrix element $M = D'\omega_q/(2\rho s^2)$ yielding a transition probability

$$W_{\mathbf{k}',\mathbf{k}} = \alpha \sum_{\mathbf{q}} |\langle \mathbf{k}'|\mathbf{k}\rangle|^2 \omega_{\mathbf{q}}^2 \Big[N^2_{\omega_{\mathbf{q}}} \delta_+ + (N_{\omega_{\mathbf{q}}}+1)^2 \delta_- \Big], \tag{3.24}$$

where $\alpha = \frac{2\pi N D'^2}{\hbar(2\rho s^2)^2}$, $\delta_\pm = \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} \pm 2\omega_{\mathbf{q}})$. Here the coherence factor is $|\langle \mathbf{k}'|\mathbf{k}\rangle|^2 = \big[1 \pm \cos(\theta_{\mathbf{k}} - \theta_{\mathbf{k}'})\big]/2$, with the plus (minus) sign for intra-band (inter-band) processes.

The two-phonon cooling power can be obtained from the transition probability by using Eq.(3.9). We find

$$\mathcal{J}_{2-\rm ph} = \big[\nu(\mu)\big]^2 \alpha \int \frac{d^2q}{(2\pi)^2} 4\omega_q^4 \Delta, \tag{3.25}$$

where $\Delta = (N_{\omega_{\mathbf{q}}}+1)^2 \tilde{N}_{2\omega_{\mathbf{q}}} - N^2_{\omega_{\mathbf{q}}}(\tilde{N}_{2\omega_{\mathbf{q}}}+1)$. We note that Δ vanishes when $T_{\rm el} = T_{\rm ph}$, satisfying detailed balance. Expanding in a small $\Delta T = T_{\rm el} - T_{\rm ph}$ and integrating over $\mathbf{q}$ we obtain the ratio $\mathcal{J}_3/\mathcal{J}_0 = (D'/D)^2 (T/T_0)^5$, where $k_{\rm B}T_0 = 0.86\,(\hbar^2\rho s^4 k_{\rm B}^2 T_{\rm BG}^2)^{1/5}$. Taking D' of order unity in atomic units, $D' \sim e^2/a \sim 10\,{\rm eV}$, we obtain a crossover temperature which for typical dopings is $5-10$ larger than our T_* estimate for disorder assisted processes. This means that the two-phonon processes can dominate only in very clean systems.

3.4.2 Sequential Two Phonon Scattering

In this section we consider a two-phonon channel that emit (or absorb) phonons sequentially. Scattering from this two step two-phonon process comes from the standard deformation coupling, described in Eqn. 3.4 and has a transition probability that vanishes. We can see this by examining the matrix element in this process described by

$$M = \langle \mathbf{k}' | M^{(0)}(\mathbf{q_1}) G^{q_2} M^{(0)}(\mathbf{q_2}) + M^{(0)}(\mathbf{q_2}) G^{q_1} M^{(0)}(\mathbf{q_1}) | \mathbf{k} \rangle. \tag{3.26}$$

where $G^{q_1} = 1/(\epsilon - H_0(\mathbf{k} - \mathbf{q_1}))$ and $G^{q_2} = 1/(\epsilon - H_0(\mathbf{k} - \mathbf{q_2}))$. Since $|\mathbf{q_1}|, |\mathbf{q_2}| \approx q_T \gg k_F$, we can neglect $\mathbf{k}, \mathbf{k}'$ and the two phonons scattered counterpropagate, i.e. $\mathbf{q_1} \approx -\mathbf{q_2} = \mathbf{q}$. This means that $G^{q_1} \approx -1/H_0(\mathbf{q})$ and $G^{q_2} \approx -1/H_0(-\mathbf{q})$. Now we note that $M^{(0)}$ is an identity in graphene's pseudospin space and hence commutes with H_0 and itself. Since $H_0(-\mathbf{q}) = -H_0(\mathbf{q})$, the matrix element and the resulting transition probability vanishes. This, in turn, means that the cooling power *vanishes*.

This vanishing result strictly applies to a pristinely undoped case (that is $\mu = 0$) and a non-vanishing result can be obtained for finite doping, $\mu \neq 0$. However, the cooling power obtained from finite doping is negligibly small.

3.4.3 Flexural Phonon Supercollisions

Finally, we analyze cooling in free standing graphene in the absence of disorder. In this case, an important contribution arises due to flexural phonons [81, 82], which contribute to the deformation tensor via $u_{ij} = 1/2(\partial_i u_j + \partial_j u_i + \partial_i h \partial_j h)$, with $\mathbf{u}$ and h the in-plane and out-of-plane displacements. Flexural modes have quadratic dispersion $\tilde{\omega}_{\mathbf{q}} = \kappa |\mathbf{q}|^2$ with $\kappa \approx 4.6 \times 10^{-7}\,\mathrm{m^2 s^{-1}}$ [81,82]. Electron-phonon coupling is described by the same deformation potential as above, Eq.(3.4).

The processes involving pairs of nearly counterpropagating flexural phonons are

analyzed as follows. Using the momentum representation, $h_{\mathbf{q}} = \sqrt{\hbar/2\rho\tilde{\omega}_q}\big(b_{\mathbf{q}} + b^{\dagger}_{-\mathbf{q}}\big)$, we consider the emission/absorption of two flexural phonons with momenta $\mathbf{q_1}$ and $\mathbf{q_2}$. For $T \gg T_{\rm BG}^{\rm flex} = \hbar\kappa k_F^2$ (for typical densities, $T_{\rm BG}^{\rm flex}$ is well below 1 K), we can set $\mathbf{q_1} \approx -\mathbf{q_2} = \mathbf{q}$, yielding the transition probability

$$W_{\mathbf{k}',\mathbf{k}}=\frac{2\pi}{\hbar}\sum_{\mathbf{q}}|\langle\mathbf{k}'|\mathbf{k}\rangle|^2M^2\Big[N^2_{\omega_{\mathbf{q}}}\delta_+ + (N_{\omega_{\mathbf{q}}}+1)^2\delta_-\Big], \tag{3.27}$$

$\delta_\pm = \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} \pm 2\tilde{\omega}_{\mathbf{q}})$. Here the matrix element is $M = \frac{D\hbar}{4\rho\kappa}$ [81,82] and the coherence factor is $|\langle\mathbf{k}'|\mathbf{k}\rangle|^2 = \big[1 \pm \cos(\theta_{\mathbf{k}} - \theta_{\mathbf{k}'})\big]/2$, with the plus (minus) sign for intra-band (inter-band) processes. This gives the energy-loss power

$$\mathcal{J}_{\rm flex} = \sum_q (2\hbar\tilde{\omega}_q)^2\left[(N_{\omega_{\mathbf{q}}}+1)^2N^{\rm el}_{2\omega_{\mathbf{q}}} - N^2_{\omega_{\mathbf{q}}}(N^{\rm el}_{2\omega_{\mathbf{q}}}+1)\right]$$

where $\sum_q ... = \frac{\pi N D^2\hbar^2}{16\rho^2\kappa^2}\big[\nu(\mu)\big]^2\int\frac{d^2q}{(2\pi)^2}...$ We note that the above expression vanishes when $T_{\rm el} = T_{\rm ph}$.

We linearize $\mathcal{J}_{\rm flex}$ in $\Delta T = T_{\rm el} - T_{\rm ph}$ to obtain

$$\mathcal{J}_{\rm flex} = A_{\rm flex}T^2\Delta T, \qquad A_{\rm flex} = 0.12\frac{ND^2\nu^2(\mu)k_{\rm B}^3}{\rho^2\kappa^3}, \tag{3.28}$$

which scales with T the same way as Eq.(3.1). Flexural phonons dominate over the one-phonon contribution at

$$T > T_*^{\rm flex} = \sqrt{\frac{B}{A_{\rm flex}}} = \Big(\frac{\pi\rho\kappa^3}{0.24\hbar s^2}\Big)^{1/2}T_{\rm BG} \approx 10\,T_{\rm BG}. \tag{3.29}$$

The value $T_*^{\rm flex}$ is considerably larger than T_* for disorder-assisted cooling estimated above. A comparison with Eq(3.1) yields $\mathcal{J}_{\rm flex}/\mathcal{J} \approx k_F\ell/200$ which is small for typical $k_F\ell$.Thus the contribution (3.28) is relatively weak under realistic conditions. For graphene on substrate this contribution is further diminished as flexural modes get pinned, gapped, and stiffened by the substrate.

3.5 Summary

In summary, graphene stands out as a unique system where disorder-limited cooling is the *leading* contribution over a wide range of temperatures, including room temperature. While first order processes such as single optical (acoustic) phonon emission are inefficient cooling channels, paradoxically higher-order processes (supercollisions) alleviate this bottleneck. There are a wide variety of supercollisions in graphene the most important of which being disorder-assisted supercollisions. As a result, varying the amount of disorder can be used as a knob to tailor and control a variety of hot carrier effects in graphene. Tuning disorder can be achieved by well established techniques, including current-annealing and using different substrates (e.g. SiO2 or BN). The characteristic T^3 dependence, Eq.(3.1), and power-law cooling dynamics, Eq.(3.14), makes this new regime easy to identify in experiments [95, 96].

Chapter 4

The Photoexcitation Cascade in Graphene

Converting light to electrical currents or voltages is a complex, multi-step process which involves photo-excited particles and holes undergoing scattering by other charge carriers and by lattice vibrations. One of the key questions in the field of optoelectronics is identifying materials in which multiple excitation generation can occur, i.e. a single absorbed photon yielding a large number of particle-hole pairs as a result of the primary photoexcited pair producing secondary pairs. Efficient multiple excitation generation relies on a combination of characteristics such as a wide band of states with a large phase space density for pair excitations, strong electron-electron scattering, and not too strong electron-phonon interaction. While graphene is by no means a unique example of a system with these properties, it is believed to fit the bill better than other materials. This has motivated an intense investigation of photoexcitation processes in graphene-based systems [25, 31, 57, 58, 69, 77, 97–102].

Graphene possesses a number of properties that distinguish it from other optoelectronic materials. One unique aspect of graphene is its truly two-dimensional structure which renders its electronic states fully exposed. Photogenerated carriers in such a system can in principle be extracted by a vertical transfer process, e.g. in a sandwich-type tunneling

structure. Vertical carrier extraction eliminates carrier loss in lateral transport between the photoexcitation region and contacts, often a limiting factor for optoelectronic response in semiconductor systems [34]. Another distinguishing trait of graphene is slow electron-lattice cooling [11, 12, 26], which leads to extended hot carrier cooling lengths reaching a few microns in clean systems even at room temperature [21, 27, 28], see also discussion in Chapter 3. Slow cooling enhances hot-carrier effects, leading to a unique photocurrent generation mechanism [21, 74]. As a result, the optoelectronic response of graphene is particularly sensitive to the details of the photo-excitation cascade.

A number of experimental techniques have been employed to track the decay dynamics of photo-excited carriers in graphene, unraveling a complex picture of competing relaxation pathways [30,31,69,77,97,102]. Recent ultrafast optical pump-terahertz probe [30] and angle resolved photoemission spectroscopy (ARPES) [103] experiments have indicated that electron-electron scattering events occur on a fast time scale, several tens of femtoseconds highlighting the crucial role interactions play in the photo-excitation cascade (and optical response) of graphene. In particular, a recent ultrafast optical pump-terahertz probe study [30] obtained detailed information on the energy absorbed by the electronic system in the cascade following a short photoexcitation pulse. It was found that this number scales linearly with i) the number of absorbed photons, as well as with ii) the energy of individual photons[1]. These results indicate that the decay of photo-excited carriers was dominated by electron-electron scattering events rather than the emission of phonons (photoexcitation cascade dominated by phonon emission is not expected to show scaling with photon energy). These results, reproduced in Fig.4.1 (b), highlight the crucial role that interactions play in the photo-excitation cascade (and optical response) of graphene.

[1]More details of this experiment can be found at the end of this chapter. Additionally Ref. [30] contains a detailed account of the experiment and analysis performed.

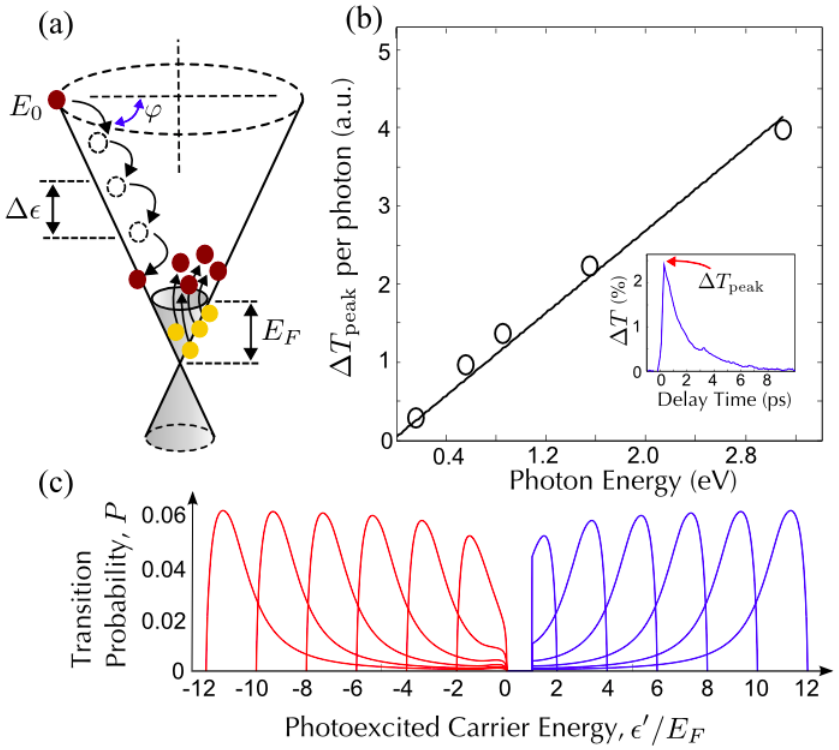

Figure 4.1: (a) Impact excitation (IE) cascade of a photo-excited carrier with initial energy E_0. Each cascade step involves electron-hole pair excitations with energy $\Delta\epsilon \sim E_F$, where E_F is Fermi energy. The net number of generated pairs and the relaxation rate depend strongly on E_F [see Eqs.(4.1),(4.14)], and thus can be tuned by gate voltage. (b) The energy captured by as a hot carrier distribution per photon measured by an optical pump-terahertz probe technique in CVD graphene. Typical experimental trace for differential transmission ΔT as a function of pump-probe delay time, shown in the inset, features a pronounced peak. The peak height is proportional to the pump induced heat captured in a hot carrier distrbution. Peak value $\Delta T_{\rm peak}$, normalized by absorbed photon density, is shown as a function of photon energy (adapted from Ref. [30]). Note linear scaling of $\Delta T_{\rm peak}$ vs. pump photon energy. (c) Transition probability (in units of $\hbar^{-1}$) obtained from Eq.(4.13). IE processes with different initial energies, $\epsilon_i/E_F = -12, -10, \ldots, 10, 12$ (here $\Delta\epsilon = \epsilon_i - \epsilon'$). Electron (hole) contributions shown by blue (red) curves.

4.1 Photo-Excited Carrier Dynamics

Despite intense interest, the photo-excitation cascade in graphene remains poorly understood. Theory predicts that the linear dispersion of charge carriers acquires a negative curvature due to electron-electron interactions, $d^2\epsilon(k)/dk^2 < 0$, which inhibits decay via electron-electron scattering in undoped graphene [98, 104]. However, while the prediction of negative curvature appears to be in agreement with transport measurements [105], ARPES experiments support the notion of interaction-mediated decay [106, 107] and pump-probe experiments point to the crucial role interactions play in the photo-excitation cascade [30]; interaction-induced quasiparticle decay remains the subject of ongoing debate [25, 98, 108, 109].

In this chapter[2], we show that the photo-excitation cascade in doped graphene is distinct from that in undoped graphene. We identify *impact excitation* (IE) as the scattering process (see intra-band carrier-carrier scattering process in Fig. 4.1(a)) that dominates carrier relaxation dynamics in this system. Multiple secondary electron-hole (e-h) pairs produced by IE scattering involving a photo-excited carrier and ambient carriers in the Fermi-sea can lead to the generation of multiple hot carriers. Our analysis predicts that IE processes result in a chain-like cascade consisting of sequential steps with relatively *small energy loss* per step $\Delta\epsilon \sim E_F$, where E_F is the Fermi energy in graphene doped away from neutrality (see Fig.4.1(a) and (c)). As we shall see, both the number of pairs produced in the cascade and the characteristic energy for the pairs are highly sensitive to doping. As a result, the key parameters of photo-excitation cascade in graphene are expected to be gate tunable in a wide range.

[2]The introduction, and the first section in this chapter is reproduced in part from JCW Song, KJ Tielrooij, FHL Koppens, LS Levitov, Photoexcited carrier dynamics and impact-excitation cascade in graphene, *Physical Review B* **87**, 155429 Copyright (2013) by the American Physical Society. Sec. 4.2 is based on work described in Ref. [30]; the text in Ref. [30] has not been reproduced.

As we argue below, the IE rate takes the highest values allowed by unitarity, $\Gamma \sim E_F/(2\pi\hbar)$ (these values are consistent with the inelastic lifetimes in Ref. [63,109,110]). This fast characteristic rate makes this scattering process a highly efficient relaxation pathway which dominates over phonon-mediated pathways in a wide range of energies. As a result, the photo-excitation cascade proceeds in the fashion depicted in Fig. 4.1(a) and (c) making both the number of pairs produced in the cascade and the characteristic energy for the pairs are highly sensitive to doping ; the key parameters of the photo-excitation cascade in graphene are expected to be gate tunable in a wide range.

The dependence on excitation energy E_0 and Fermi energy E_F provides clear experimental signatures of this relaxation mechanism. In particular, the average number of e-h pairs produced in the cascade triggered by a *single photo-excited electron*, $\langle N\rangle$, is

$$\langle N\rangle = \int_{E_L}^{E_0} \frac{d\epsilon}{\langle\Delta\epsilon\rangle}, \quad \langle\Delta\epsilon\rangle = \frac{\mathcal{J}_{\rm el}(\epsilon)}{\Gamma(\epsilon)}, \tag{4.1}$$

where $\langle\Delta\epsilon\rangle$ is the average energy loss per step, $\mathcal{J}_{\rm el}(\epsilon)$ and $\Gamma(\epsilon)$ are the IE energy-relaxation and scattering rates respectively (see Eq.(4.14)). Here $E_L \approx E_F$ is a low-energy cutoff corresponding to the energy below which IE processes are quenched; we use $E_L = 2E_F$ (see discussion below). Fig.4.2(a) indicates that $\langle N\rangle$ exceeds unity and grows quickly for E_0 above few E_F (red curve). Since $\langle\Delta\epsilon\rangle \sim E_F$, we find that $\langle N\rangle$ scales as E_0/E_F. In particular, an approximately linear dependence $\langle N\rangle \approx 0.55E_0/E_F$ is found for $E_0/E_F \gg 1$.

Similarly, the time it takes for the photo-excited electron to completely decay, Δt, also exhibits strong E_0 and E_F dependence. This fairly short time, on the order of hundreds of femtoseconds, is

$$\Delta t = \int_{E_L}^{E_0} \frac{d\epsilon}{\mathcal{J}_{\rm el}(\epsilon)} = \frac{\mathcal{G}(E_0/E_F)}{E_F[\text{eV}]}\,\text{fs}, \tag{4.2}$$

where $\mathcal{G}$ is a dimensionless scaling function (blue curve in Fig. 4.2(a)). As shown in Fig. 4.2(a), $\mathcal{G}$ scales approximately linearly with E_0/E_F, yielding a Δt that scales linearly with

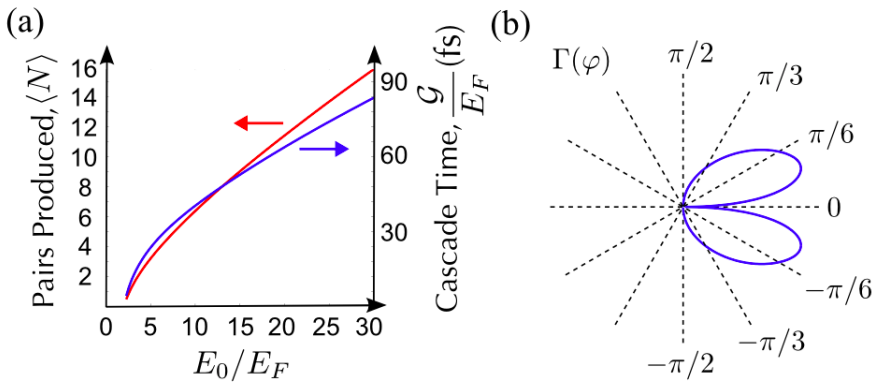

Figure 4.2: (a) Average net number of e-h pairs produced in the cascade triggered by a photo-excited electron with energy E_0 (red curve). Cascade duration, $\Delta t = \mathcal{G}/(\mu\,[\text{eV}])\text{fs}$, see Eq.(4.2) (blue curve). (b) Angular dependence for the e-h excitation rate $\Gamma(\varphi)$, where φ is the angle between $\mathbf{k_1}$ and $\mathbf{q}$, see Fig.4.3(a).

the excitation energy. For a typical doping value of $E_F = 0.2\,\text{eV}$ and initial photo-excited carrier energy $E_0 = 1\,\text{eV}$ we find $\Delta t \approx 0.12\,\text{ps}$, far faster than typical electron-lattice cooling time scales found in graphene [11, 12, 26]. This separation of time scales means that the energy relaxation cascade occurs independently of electron-lattice cooling.

Lastly, the angular distribution for impact excitation transitions is highly anisotropic. This produces a strong search-light-type structure peaked along the preferred direction of momentum transfer shown in Fig. 4.2 (b).

4.1.1 Impact Excitation Scattering

Our system is described by the Hamiltonian for $N = 4$ species of massless Dirac particles,

$$\mathcal{H} = \sum_{\mathbf{k},i} \psi^\dagger_{\mathbf{k},i}(v\sigma\cdot\mathbf{k})\psi_{\mathbf{k},i} + \mathcal{H}_{\text{el-el}}, \tag{4.3}$$

$$\mathcal{H}_{\text{el-el}} = \frac{1}{2}\sum_{\mathbf{q},\mathbf{k},\mathbf{k'},i,j} V(\mathbf{q})\psi^\dagger_{\mathbf{k}+\mathbf{q},i}\psi^\dagger_{\mathbf{k'}-\mathbf{q},j}\psi_{\mathbf{k'},j}\psi_{\mathbf{k},i}. \tag{4.4}$$

Here $i, j = 1...N$ and $V(\mathbf{q}) = 2\pi e^2/|\mathbf{q}|\kappa$ is the Coulomb interaction. Importantly, transitions in a massless Dirac band governed by the Hamiltonian (4.3) are subject to certain kinematical constraints [25,98]. These constraints arise due to the combined effect of linear dispersion

in two Dirac cones, $E_{\pm}(\mathbf{p}) = \pm v|\mathbf{p}|$, and momentum conserving character of carrier scattering. Here we analyze the simplest case of a two-body collision. Each of the two particles participating in a collision can make transitions between states in the upper and lower Dirac cones which we denote by + and −. Two kinds of transitions can be distinguished: intra-band transitions ($+ \to +$ or $- \to -$) and inter-band transitions ($+ \to -$ or $- \to +$). Since momentum change in any transition satisfies $||\mathbf{p}_1| - |\mathbf{p}_2|| < |\mathbf{p}_1 - \mathbf{p}_2| < |\mathbf{p}_1| + |\mathbf{p}_2|$, the intra-band transitions can only occur when the energy and momentum change are related by $|\Delta\epsilon| \leq v|\Delta p|$, whereas the inter-band transitions are possible only when $|\Delta\epsilon| \geq v|\Delta p|$.

The scattering process of interest, pictured in Fig.4.3(a), involves a photo-excited carrier with high energy and momentum $\epsilon_{\mathbf{k_1}} \gg E_F$, $|\mathbf{k}_1| \gg k_F$, which is scattered to a lower energy state having momentum $\mathbf{k}_1'$ with recoil momentum $\mathbf{q} = \mathbf{k_1} - \mathbf{k}_1'$ given to an electron in the Fermi sea. The latter process results in a particle-hole pair excitation, as depicted by a transition from $\mathbf{k_2}$ to $\mathbf{k}_2'$ in Fig. 4.3 (a). The transition rate for this process, evaluated by the standard Golden Rule approach, takes the form

$$\begin{aligned} W_{\mathbf{k}_1',\mathbf{k_1}} &= \frac{2\pi N}{\hbar} \sum_{\mathbf{q},\mathbf{k_2},\mathbf{k}_2'} f_{\mathbf{k_2}}(1 - f_{\mathbf{k}_2'}) F_{\mathbf{k_2},\mathbf{k}_2'} |\tilde{V}_{\mathbf{q}}|^2 \\ &\quad \times \delta_{\mathbf{k}_1',\mathbf{k_1}+\mathbf{q}} \delta_{\mathbf{k}_2',\mathbf{k_2}-\mathbf{q}} \delta(\epsilon_{\mathbf{k}_1'} - \epsilon_{\mathbf{k_1}} + \epsilon_{\mathbf{k}_2'} - \epsilon_{\mathbf{k_2}}). \end{aligned} \tag{4.5}$$

Here $f_{\mathbf{k}}$ is a Fermi function, and $F_{\mathbf{k},\mathbf{k}'} = |\langle \mathbf{k}'s'|\mathbf{k}s\rangle|^2$ is the coherence factor ($s, s' = \pm$ label states in the electron and hole Dirac cones). We treat the Coulomb interaction which mediates scattering between the photo-excited carrier and the carriers in the Fermi sea by accounting for dynamical screening in the RPA approximation:

$$\tilde{V}_{\mathbf{q}} = \frac{V_{\mathbf{q}}^0}{\varepsilon(\omega, \mathbf{q})}, \quad \varepsilon(\omega, \mathbf{q}) = 1 - V_{\mathbf{q}}^0 \Pi(\mathbf{q}, \omega), \tag{4.6}$$

where $V_{\mathbf{q}}^0 = 2\pi e^2/|\mathbf{q}|\kappa$ and $\varepsilon(\omega, \mathbf{q})$ describes dynamical screening. Here Π is the polarization

operator

$$\Pi(\mathbf{q},\omega) = N\sum_{\mathbf{k},s,s'} F_{\mathbf{k},\mathbf{k}+\mathbf{q};ss'}\frac{f(\epsilon_{\mathbf{k},s}) - f(\epsilon_{\mathbf{k}+\mathbf{q},s'})}{\omega + \epsilon_{\mathbf{k},s} - \epsilon_{\mathbf{k}+\mathbf{q},s'} + i0}, \tag{4.7}$$

with band indices $s, s' = \pm$. This includes both intra- ($s = s'$) and inter- ($s \neq s'$) band contributions [111].

For Eq.(4.5) to give a non-vanishing result, the transitions $\mathbf{k}_1 \to \mathbf{k}_1'$, $\mathbf{k}_2 \to \mathbf{k}_2'$ must occur in like pairs, both intra-band or both inter-band. Since $\mathbf{k}_1 \to \mathbf{k}_1'$ is restricted to be within a single band, $\mathbf{k}_2 \to \mathbf{k}_2'$ must also be intra-band. As a result, relaxation via inter-band scattering is blocked, whereas intra-band scattering is allowed. Kinematical blocking of inter-band processes can in principle be relieved by three-body (or higher-order) collisions (not discussed here). Such processes may become important at strong excitation, however they are expected to be weak in the low excitation power regime.

As shown below, the typical energy of an excited pair is much smaller than the photo-excitation energy $\epsilon_{\mathbf{k}_1}$. Anticipating this result, it is convenient to factorize the transition rate by expressing it through the spectrum of secondary pair excitations [112]. This can be accomplished by writing $\delta(\epsilon_{\mathbf{k_1'}} - \epsilon_{\mathbf{k_1}} + \epsilon_{\mathbf{k_2'}} - \epsilon_{\mathbf{k_2}}) = \int_{-\infty}^{\infty} d\omega\delta(\epsilon_{\mathbf{k_1'}} - \epsilon_{\mathbf{k_1}} + \omega)\delta(\epsilon_{\mathbf{k_2'}} - \epsilon_{\mathbf{k_2}} - \omega)$. Next, we use the identity $f_{\mathbf{k_2}}(1 - f_{\mathbf{k_2'}}) = (f_{\mathbf{k_2}} - f_{\mathbf{k_2'}}) \times \big(N(\epsilon_{\mathbf{k_2'}} - \epsilon_{\mathbf{k_2}}) + 1\big)$, where $N(\omega) = 1/(e^{\omega/k_BT} - 1)$ is the Bose function taken at the electron temperature. Lastly, we express the sum of $(f_{\mathbf{k_2}} - f_{\mathbf{k_2'}})\delta(\epsilon_{\mathbf{k_2'}} - \epsilon_{\mathbf{k_2}} - \omega)$ through a suitably defined susceptibility

$$\chi''(\mathbf{q},\omega) = N\sum_{\mathbf{k}} F_{\mathbf{k},\mathbf{k}+\mathbf{q}}(f_{\mathbf{k}} - f_{\mathbf{k}+\mathbf{q}})\delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \omega), \tag{4.8}$$

which can also be written as $\chi''(\mathbf{q},\omega) = -\frac{1}{\pi}\operatorname{Im}\Pi(\mathbf{q},\omega)$. [3] This yields a compact and

[3]The "factorization" process described above is particularly useful since $\chi''(\mathbf{q},\omega)$ can be written in terms of the imaginary part of graphene's polarization function that is well known. For our subsequent calculation, we used the polarization function by Hwang and Das Sarma in [111]. Since only intraband processes are allowed in IE, we used only the part of the polarization function intraband processes to obtain $\chi''(\mathbf{q},\omega)$.

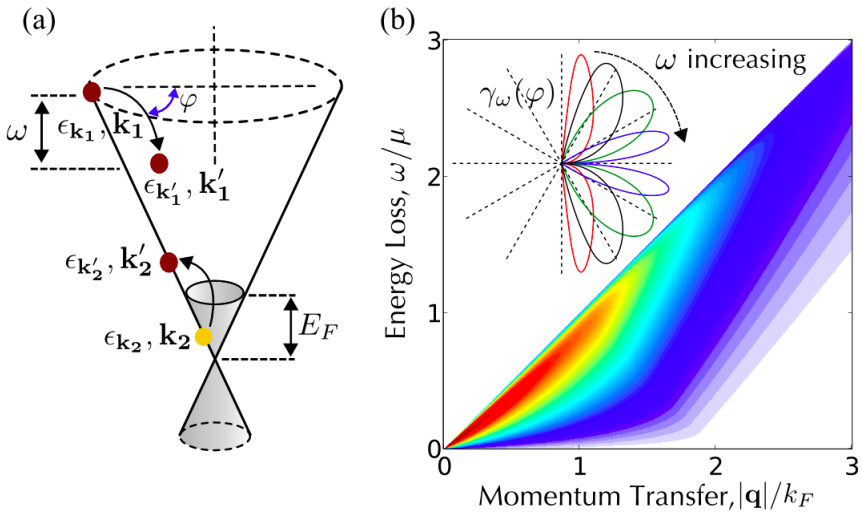

Figure 4.3: (a) Kinematics of intraband carrier-carrier scattering in doped graphene. A photo-excited electron makes transition from $\mathbf{k_1}$ to $\mathbf{k}'_\mathbf{1}$ by exciting an electron-hole pair from the Fermi sea from $\mathbf{k_2}$ to $\mathbf{k}'_\mathbf{2}$. (b) Spectral function $R(\mathbf{q},\omega)$ of particle-hole excitation as a function of momentum transfer and energy transfer per scattering event. (inset) Angular distribution of normalized energy resolved transition rate (see Eq. 4.12, for fixed values of $\omega/E_F = 0.2, 0.5, 1, 10$ (red, black, green, and blue curves respectively). The near-collinear character of scattering at high ω is manifested in narrowing of the angular distribution.

intuitive expression for the total scattering rate:

$$\Gamma = \sum_{\mathbf{k}'_1} W_{\mathbf{k}'_1,\mathbf{k}_1}(1-f_{\mathbf{k}'_1})F_{\mathbf{k_1},\mathbf{k}'_1} = \int_{-\infty}^{\infty} d\omega P(\omega), \tag{4.9}$$

$$P(\omega) = A\sum_{\mathbf{q}} |\tilde{V}_\mathbf{q}|^2 F_{\mathbf{k_1},\mathbf{k}'_1}\chi''(\mathbf{q},\omega)\delta(\epsilon_{\mathbf{k}'_1} - \epsilon_{\mathbf{k}_1} + \omega), \tag{4.10}$$

where $A = \frac{2\pi}{\hbar}[N(\omega)+1)][1-f(\epsilon_{\mathbf{k_1}}-\omega)]$ and $\mathbf{k}'_1 = \mathbf{k}_1 - \mathbf{q}$.

As we show below, the typical energy and momentum transferred per scattering is of order E_F and E_F/v, respectively. These values are much smaller than those of the photo-excited electron. We can therefore approximate $F_{\mathbf{k}_1,\mathbf{k}'_1} \approx 1$, $f(\epsilon_{\mathbf{k_1}}-\omega) \approx 0$ and write the delta function as $\delta(\epsilon_{\mathbf{k}'_1} - \epsilon_{\mathbf{k}_1} + \omega) \approx \delta(v|\mathbf{q}|\cos\varphi - \omega)$, where φ is the angle between $\mathbf{k}_1$ and $\mathbf{q}$. The approximation $|\mathbf{q}| \ll |\mathbf{k}_1|$, $|\omega| \ll v|\mathbf{k}_1|$ is appropriate under realistic conditions: for example, visible light frequencies translate to $\epsilon_\mathbf{k} = hf/2 = 750\,\text{meV}$ which is considerably larger than E_F for typical doping values. Eq.(4.9) then yields the angle dependent transition

rate

$$\Gamma(\varphi) = \int_{-\infty}^{\infty} d\omega \gamma_\omega(\varphi), \tag{4.11}$$

$$\gamma_\omega(\varphi) = \int \frac{d^2q}{(2\pi)^2} R(\mathbf{q},\omega)\delta(v|\mathbf{q}|\cos\varphi - \omega), \tag{4.12}$$

where $R(\mathbf{q},\omega) = A|\tilde{V}_\mathbf{q}|^2\chi''(\mathbf{q},\omega)$. We evaluate (numerically) the spectral function $R(\mathbf{q},\omega)$ using the RPA-screened interaction, Eq.(4.6),and susceptibility expressed through the polarization function from Ref. [111] and the interaction parameter $\alpha = e^2/(\kappa\hbar v) = 0.73$. The angular distribution, $\Gamma(\varphi)$, as well as the energy resolved distribution, $\gamma_\omega(\varphi)$, feature interesting angular patterns (see Fig.4.2 (b) and Fig.4.3 inset). Note in particular a sharp search-light-type structure corresponding the preferred direction of momentum transfer, $\mathbf{q}$, in the IE process. The peaks move closer to $\varphi = 0$ as ω increases, indicating that carrier-carrier scattering with high energy transfer is nearly collinear. This is analogous to radiation pattern for an ultra-relativistic particle becoming focused along particle velocity [113].

The same approach can be used to obtain the energy spectrum of pair excitations. In the following, however, we study the *full energy dependence* of $P(\Delta\epsilon)$ not limiting ourselves to the asymptotic behavior at high photo-excited energies. Using a Jacobian to convert the delta function in energy to a delta function in angles in Eq.(4.10), we perform the angular integral in Eq.(4.10) to obtain

$$P(\Delta\epsilon) = \int_0^\infty \frac{2|\mathbf{k_1}| - (\Delta\epsilon/v\hbar) - |\mathbf{q}|\cos\varphi}{(2\pi)^2|\mathbf{k_1}||\mathbf{q}|\hbar v\sin\varphi} R(\mathbf{q},\Delta\epsilon) q dq, \tag{4.13}$$

where φ is the angle between $\mathbf{k_1}$ and $\mathbf{q}$ and satisfies $(|\mathbf{k_1}|^2 - 2|\mathbf{k_1}||\mathbf{q}|\cos\varphi + |\mathbf{q}|^2)^{1/2} - |\mathbf{k_1}| = \Delta\epsilon/(v\hbar)$. Numerically integrating Eq.(4.13), using Π from Ref. [111], and taking $\epsilon \gg k_B T$ yields transition probabilities $P(\Delta\epsilon)$ shown in Fig. 4.1(c) for different initial photo-excited energies $\epsilon_i = \epsilon_{\mathbf{k}_1}$. We find that $P(\Delta\epsilon)$ peaks close to $\Delta\epsilon \approx E_F$ and decays rapidly for $\Delta\epsilon \gg E_F$. This non-monotonic dependence arises from the competition between

the available phase space, which grows with $\Delta\epsilon$, and the Coulomb Interaction form factor, which decreases with $|\mathbf{q}|$.

The efficiency of IE scattering can be linked to the large values of E_F in graphene. The relation between efficiency and E_F can be clarified by simple dimensional analysis. We note that $P(\Delta\epsilon)$ depends on $\Delta\epsilon$ essentially via the dimensionless parameter $x = \Delta\epsilon/E_F$. This is clearly seen e.g. from pair excitation spectrum shown for different values of initial energy $\epsilon_i = \epsilon_{\mathbf{k}_1}$ in Fig.4.1(c): the width and profile of $P(\Delta\epsilon)$ has a very weak dependence on ϵ_i. This can be captured by writing the scattering rate Γ (Eq.(4.9)) as well as energy relaxation rate $\mathcal{J}_{\rm el} = \int_{-\infty}^{\infty} \Delta\epsilon P(\Delta\epsilon) d\Delta\epsilon$ in the form

$$\Gamma(\epsilon) = \frac{E_F}{\hbar}\int_0^{\epsilon/E_F} \tilde{P}(x)dx, \quad \mathcal{J}_{\rm el}(\epsilon) = \frac{E_F^2}{\hbar}\int_0^{\epsilon/E_F} x\tilde{P}(x)dx, \tag{4.14}$$

where we introduced dimensionless $\tilde{P}(x) = \hbar P(\Delta\epsilon)$.

The E_F dependences in Eq.(4.14) manifests in observables such as the average number of secondary e-h pairs produced in a single photo-excitation cascade, $\langle N\rangle$ and its total cascade time, Δt. These quantities are related via $\langle N\rangle = \int_0^{\Delta t}\Gamma dt$. Using $d\epsilon/dt = -\mathcal{J}_{\rm el}(\epsilon)$ combined with Eq.(4.14) we obtain Eq.(4.1) for $\langle N\rangle$ and Eq.(4.2) for Δt. In both cases, we used a low energy cutoff for the energy below which IE processes are quenched, $E_L \approx E_F$. Below the energy E_L, the relaxation and scattering of the carrier from impact excitation slows dramatically and other relaxation processes dominate, for example energy relaxation via the emission of acoustic phonons. In evaluating Eqs.(4.1),(4.2), we used the value $E_L = 2E_F$ below which the predicted value Δt rapidly increases. The scaling of Δt and $\langle N\rangle$ with both excitation energy E_0 and doping in Fig. 4.2(a) are clear experimental signatures of IE. Currently, particle-hole pair production and cascade times are the subject of intense experimental interest [30, 59, 114].

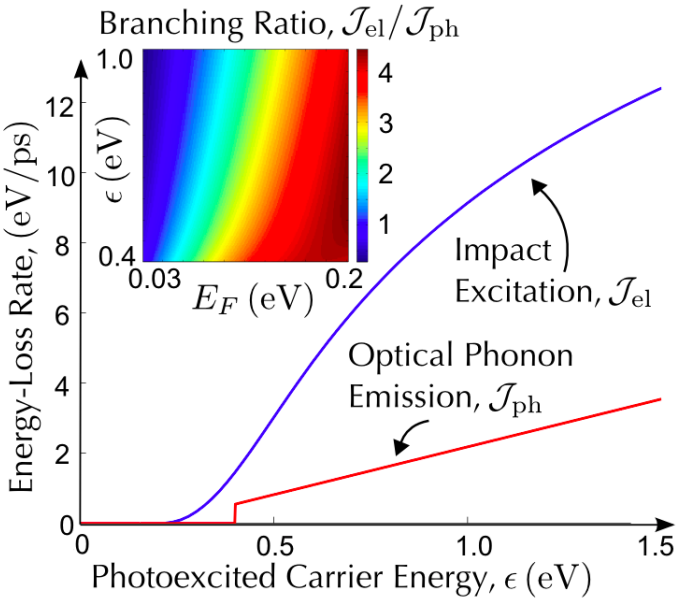

Figure 4.4: Energy-loss rate via impact excitation, $\mathcal{J}_{\rm el}$ [blue curve, see Eq.(4.14)], and optical phonon emission, $\mathcal{J}_{\rm ph}$ [red curve, see Eq.(4.19)] for a typical doping of $E_F = 0.2\,\text{eV}$. (Inset) Branching ratio, $\mathcal{J}_{\rm el}/\mathcal{J}_{\rm ph}$ vs. ϵ and E_F.

4.1.2 Optical Phonon Emission

Here we compare energy relaxation from IE processes with the contribution of other potentially significant channels. In particular, electron-phonon scattering leads to a direct transfer of energy to the lattice degrees of freedom without creation of secondary electron-hole excitations. We focus on the contribution of optical phonons, which under normal circumstances is more important than that of acoustic phonons. Our estimate shows that under realistic conditions the contribution of optical phonons to the energy relaxation rate is weaker than that due to carrier-carrier scattering, $\mathcal{J}_{\rm ph} \lesssim \mathcal{J}_{\rm el}$. The transition rate due to electron-phonon scattering can be described by Fermi's golden rule

$$W^{\rm el-ph}_{\mathbf{k}',\mathbf{k}} = \frac{2\pi N}{\hbar}\sum_{\mathbf{q}} |M(\mathbf{k}',\mathbf{k})|^2 \delta\big(\Delta\epsilon_{\mathbf{k}',\mathbf{k}} + \omega_{\mathbf{q}}\big) \times \delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}}(N(\omega_{\mathbf{q}})+1), \quad \Delta\epsilon_{\mathbf{k}',\mathbf{k}} = \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}, \tag{4.15}$$

where $\omega_{\mathbf{q}} = \omega_0 = 200\,\text{meV}$ is the optical phonon dispersion relation, and $N(\omega_{\mathbf{q}})$ is a Bose function. Here $\mathbf{k}$ is the initial momentum of the photo-excited electron, $\mathbf{k}'$ is the momentum

it gets scattered into, and $\mathbf{q}$ is the momentum of the optical phonon.The electron-phonon matrix element $M(\mathbf{k}', \mathbf{k})$ is

$$|M(\mathbf{k}', \mathbf{k})|^2 = g_0^2 F_{\mathbf{k},\mathbf{k}'}, \quad g_0 = \frac{2\hbar^2 v}{\sqrt{\rho\omega_0 a^4}}, \tag{4.16}$$

where $F_{\mathbf{k},\mathbf{k}'}$ is the coherence factor for graphene, g_0 is the electron-optical phonon coupling constant [11, 12, 50], ρ is graphene's mass density, and $a = 0.142\,\text{nm}$ is the distance between nearest neighbor carbon atoms. The energy-loss rate of the photo-excited carrier at energy ϵ due to the emission of an optical phonon is

$$\mathcal{J}_{\text{ph}}(\epsilon) = \sum_{\mathbf{k}'} W^{\text{el-ph}}_{\mathbf{k}',\mathbf{k}}(\epsilon'_k - \epsilon)\big[1 - f(\epsilon_{\mathbf{k}'})\big]. \tag{4.17}$$

Integrating over $\mathbf{q}$ and $\mathbf{k}'$ we obtain

$$\mathcal{J}_{\text{ph}}(\epsilon) = \frac{\pi N}{\hbar}\omega_0 g_0^2\big[1 - f(\epsilon - \omega_0)\big](N(\omega_0) + 1)\nu(\epsilon - \omega_0), \tag{4.18}$$

where $\nu(\epsilon) = \epsilon/(2\pi v^2\hbar^2)$ is the electron density of states in graphene. Hence, $\mathcal{J}_{\text{ph}}(\epsilon)$ varies linearly with the photo-excited carrier energy $\epsilon > \omega_0$ and vanishes for $\epsilon < \omega_0$. Because the electron-phonon coupling with optical phonon is a constant, this result is to be expected from the increased phase space to scatter into at higher photo-excited carrier energy.

To get an estimate of the energy relaxation rate, we estimate $(N(\omega_0) + 1) \approx 1$ and $1 - f(\epsilon - \omega_0) \approx \Theta(\epsilon - \omega_0 - E_F)$ to obtain

$$\mathcal{J}_{\text{ph}}(\epsilon) \approx \frac{\epsilon - \omega_0}{\tau_0}\Theta(\epsilon - \omega_0 - E_F), \quad \tau_0 = \frac{2v^2\hbar^3}{N\omega_0 g_0^2} \tag{4.19}$$

Using $\rho = 7.6 \times 10^{-11}\,\text{kg cm}^{-2}$, we find $\tau_0 \approx 350\,\text{fs}$.

Using $P(\Delta\epsilon)$ evaluated from Eq.(4.13), we can compare the IE energy relaxation rate $\mathcal{J}_{\text{el}}$ with the energy-loss rate due to optical phonons $\mathcal{J}_{\text{ph}}$. As illustrated in Fig. 4.4, for a typical doping of $E_F = 0.2\,\text{eV}$, the rate $\mathcal{J}_{\text{el}}$ overwhelms the rate $\mathcal{J}_{\text{ph}}$ over the entire spectrum of photo-excited carrier energies. We also analyzed the branching ratio $\mathcal{J}_{\text{el}}/\mathcal{J}_{\text{ph}}$

shown in the inset of Fig. 4.4 as a function of carrier density and photo-excited carrier energy. Interestingly, the density dependence of $\mathcal{J}_{\rm el}$ translates into gate-tunable branching ratio $\mathcal{J}_{\rm el}/\mathcal{J}_{\rm ph}$.

4.2 Photoexcitation Cascade

The picture of fast carrier-carrier scattering in this chapter supplements slower electron-lattice cooling discussed in the previous chapter to give a full description of how energy relaxation occurs from initial photoexcitation to final relaxation back to equilibrium. Photexcitation in doped graphene proceeds as depicted in Fig. 4.5: Light with energy, $hf > 2E_F$, impinging on graphene creates high energy electrons and holes far above the fermi surface forming a dileneated distribution of carriers [69, 103] separate from the ambient Fermi distribution of carriers (small grey peak). The separate distribution of carriers at high energy then relax, losing energy to phonons or scattering with ambient carriers (Auger processes). As we saw in the previous section, the scattering off ambient carriers far out paces scattering off phonons. This fast thermalization process is critical because it determines the amount of energy captured by the electronic system. After thermalization, a hot carrier distribution is formed with an elevated electronic temperature depicted in the second panel of Fig. 4.5. This hot carrier distribution subsequently cools down when hot carriers close to the Fermi surface emit phonons losing energy to the lattice (as described in the previous chapter). This picture of the photoexcitation cascade is borne out in numerous experiments [30, 66, 69, 77, 103, 115, 116].

The efficient IE scattering means that the electronic system is very good at absorbing energy from high energy photoexcited electrons; during the thermalization process IE overwhelms optical phonon emission. This can be quantified by examining the amount

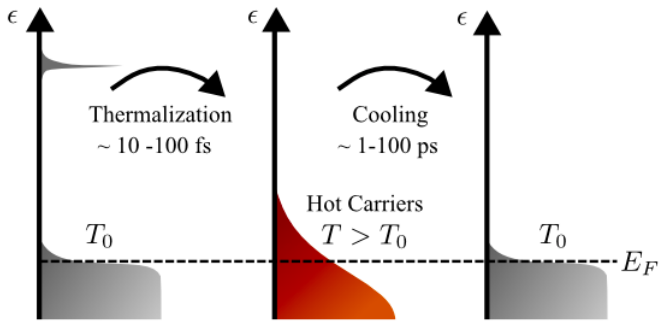

Figure 4.5: Energy relaxation of photoexcited carrier distribution occurs in two stages. First, fast thermalization (tens to hundreds of femtoseconds) of a high energy distribution occurs in which Auger processes and phonon emission takes place relaxing it down to low energies (left panel). A hot carrier distribution is then set up (middle panel). Second, electron-lattice cooling takes place over longer time scales (several to a hundred picoseconds) relaxing the hot carrier distribution back to equilibrium (right panel).

of heat absorbed, ΔQ_{el}, by the electronic system from the cascade of *a single photoexcited carrier* at $\epsilon = hf/2$

$$\Delta Q_{\text{el}} = \int_0^{t_0} \mathcal{J}_{\text{el}} dt = \int_{E_F}^{hf/2} \frac{d\epsilon}{1 + \mathcal{J}_{\text{ph}}/\mathcal{J}_{\text{el}}}, \quad \frac{d\epsilon}{dt} = -\Big(\mathcal{J}_{\text{el}} + \mathcal{J}_{\text{ph}}\Big) \tag{4.20}$$

For large enough $E_F \gtrsim 0.1$ (see Fig 4.4), $\mathcal{J}_{\text{ph}}/\mathcal{J}_{\text{el}} \ll 1$. As a result, a large fraction of the energy from the photoexcited carriers is absorbed as heat in the electronic system. This precipitates an increase in the electronic temperature, T_{el}, and generates hot carriers which can drive currents in an optoelectronic system (for example via the photocurrent response discussed in Chapter 2).

4.2.1 Measuring Hot Carrier Temperature

Measuring the electronic temperature is an ideal way of experimentally observing energy flows in graphene. Since fast carrier-carrier scattering around the Fermi surface (tens of femtoseconds) allows for an electronic temperature to be established quickly, the electronic temperature can be used to characterize both the short timescale thermalization and longer

timescale cooling. Indeed, tracking the temperature dynamics of hot carriers [27] provide a sensitive probe of the cooling mechanisms as discussed below. There are a variety of ways of measuring electronic temperature in graphene, including photocurrent [27], noise-thermometry [28], angle resolved photoemission [103], and THz conductivity [30, 115–117]. The last two have the highest temporal resolution and can be used to probe the initial thermalization dynamics of the photoexcitation cascade in pump-probe type experiments [30, 103]. Photoconductivity in pump-probe type experiments have recently received wide attention because of their ability to probe a wide variety of dynamical processes that can occur on short time scales [30, 31, 69, 77, 97, 102].

As we will see, the optical conductivity depends directly on the energy-dependent carrier distribution, $n_{\mathbf{k}}$, and gives a different value according to how far away from equilibrium the carrier distribution is. Analysis of a kinetic equation using a relaxation time approximation produces the conductivity

$$\sigma(\omega) = -\sum_{\mathbf{k}} \frac{\tau_{\mathrm{tr}}(\epsilon_{\mathbf{k}})}{1 + [\omega\tau_{\mathrm{tr}}(\epsilon_{\mathbf{k}})]^2} e^2 v_{\mathbf{k}} \nabla_{\mathbf{k}} n_{\mathbf{k}} \tag{4.21}$$

where $\tau_{\mathrm{tr}}(\mathbf{k})$ is the transport scattering time and $n_{\mathbf{k}} = (1 + e^{\beta(\epsilon_{\mathbf{k}} - \mu)})^{-1}$ depends on the electronic temperature through $\beta = 1/k_B T_{\mathrm{el}}$. Here we have shown only the real part of the conductivity since this is the most commonly measured response in most THz photoconductivity experiments [30, 115, 117]; the imaginary part can also be similarly obtained. In pump-probe type experiments, the pump-induced change in conductivity (also called photoconductivity) provides information about how far the carrier distribution is pushed out of equilibrium. As such, we will be most interested in the temperature dependent changes in conductivity $\Delta\sigma_\omega = \sigma_\omega(T_{\mathrm{el},1}) - \sigma_\omega(T_{\mathrm{el},0})$. Where $T_{\mathrm{el},0}$ denotes the initial electronic temperature and $T_{\mathrm{el},1}$ denotes the electronic temperature after pump.

Here we will be primarily interested in doped graphene since it is in this case that

IE scattering described above is most efficient. In the degenerate limit, $\mu \gg k_B T_{\rm el}$, we can employ the Sommerfeld expansion in the analysis of Eq. 4.21 yielding

$$\begin{aligned} \sigma_\omega(T) \quad &\approx \sigma_\omega(T=0) + \frac{\pi^2}{6}\nu(E_F)k_B^2T^2\frac{\partial^2 F(\epsilon)}{\partial\epsilon^2}\Big|_{\epsilon=E_F}, \\ F(\epsilon) \quad &= e^2v^2\frac{\tau_{\rm tr}(\epsilon)}{1+\omega^2[\tau_{\rm tr}(\epsilon)]^2} \end{aligned} \tag{4.22}$$

where E_F is the Fermi energy. Importantly, since the carrier density is fixed we have kept it constant by accounting for changes in chemical potential as a function of temperature $\mu \approx E_F - \frac{\pi^2}{6}k_B^2T^2/E_F$. We have also neglected the temperature dependence of the transport time, $\tau_{\rm tr}(\epsilon)$, which can be included in a more elaborate analysis. Both of these assumptions are valid when $\mu \gg k_B T_{\rm el}$. We can express the change in optical conductivity as temperature is changed, $\Delta\sigma_\omega = \sigma_\omega(T_{\rm el,1}) - \sigma_\omega(T_{\rm el,0})$, as

$$\Delta\sigma_\omega = \left(k_B^2T_{\rm el,1}^2 - k_B^2T_{\rm el,0}^2\right)\frac{\pi^2}{6}\nu(E_F)\frac{\partial^2 F(\epsilon)}{\partial\epsilon^2}\Big|_{\epsilon=E_F} \tag{4.23}$$

For graphene, $\tau_{\rm tr} \propto \epsilon$ [91, 118, 119]. As a result, $\frac{\partial^2 F(\epsilon)}{\partial\epsilon^2} < 0$ for small frequencies $\omega \sim$ THz giving a *negative* $\Delta\sigma_\omega < 0$ when electronic system gets hotter $T_{\rm el,1} > T_{\rm el,0}$. We note parenthetically, that since the heat capacity in the degenerate limit goes as $C_{\rm el} \propto T$, the absolute change in optical conductivity $|\Delta\sigma_\omega| \propto \Delta Q_{\rm el}$ making the optical conductivity a good probe of the amount of heat injected into the electronic system. In the context of probing the photoexcitation cascade, $\Delta\sigma_\omega$ is sensitive to the amount heat captured by the ambient carriers [4].

Indeed, in optical pump-terahertz probe measurements of doped graphene [30, 115–117] they found $\Delta\sigma < 0$ that was negative, agreeing with the above expectation that

[4]While we only describe the effect of an increase in electron temperature on $\Delta\sigma$, a number of other scattering channels can also affect $\Delta\sigma$. For instance, optical phonons emitted in the photoexcitation cascade can scatter with ambient carriers to change the optical conductivity measured. An estimate in Ref. [30] found that this effect was small as compared with the hot carrier effect described above for low pump fluences. However, what happens at high pump fluences remains a topic of current research.

increasing the electronic temperature decreases the optical conductivity for THz frequencies. Furthermore, in an experiment-theory collaboration with the Koppens' group in Ref. [30], the pump induced change in transmission through the graphene sheet, $\Delta T \propto -\Delta\sigma$ was analyzed as a function of pump photon energy, hf and abosorbed pump photon density, N_{photon}. It was found that ΔT scaled approximately linearly with hf and also scaled linearly with N_{photon}. The latter scaling is expected, since the having more absorbed photons clearly means that more energy is being pumped into the electronic system and $\Delta T \propto \Delta\sigma \propto \Delta Q_{\text{el}}$. The former, however, is surprising and suggests that doubling photon energy produces approximately the same effect (ie. same temperature and heat absorbed in the electron system) as doubling the number of photons.

This can be understood in simple terms by considering the heat absorbed by the electronic system in Eq. 4.20. If $J_{\text{ph}}/J_{\text{el}}$ is small, the amount of heat absorbed by the ambient electrons from the cascade of a high energy carrier scales approximately linearly, yielding a $\Delta T \propto hf$ [30]. A more sophisticated analysis can be carried out which takes into account the energy dependence of $J_{\text{ph}}/J_{\text{el}}$. Adopting such an analysis, Ref. [30] inferred a branching ratio $J_{\text{ph}}/J_{\text{el}} \sim 0.5$ for their highest energy photoexcited carriers ($hf/2 = 1.6\,\text{eV}$) and $J_{\text{ph}}/J_{\text{el}} \sim 0.25$ for their lowest energy photoexcited carriers ($hf/2 = 0.4\,\text{eV}$). Integrating this number to estimate the fraction of energy absorbed by the ambient carriers after the photoexcitation cascade yielded an estimate of $0.6 - 0.8$ [30].

While this is expected from the theoretical analysis presented above, it is remarkable given the context of the photoexcitation cascade of traditional semiconductors where a majority of the energy of the photoexcited carriers above the band edge is lost to phonons [18]. This energy loss is one of the main contributors to the Shockley-Quiesser Limit [2] that states that the maximum efficiency single junction solar cells is around 30%. The photoexcitation cascade in graphene circumvents this loss and may allow for a new

paradigm of light to electrical current conversion [34]. We note that while the electron-lattice cooling bottleneck is the main reason for the importance of hot carriers in graphene's optoeletronic response as discussed in Chapter 2, the ease with which energy is absorbed by the electronic system (embodied by the IE cascade described in this chapter) also plays an important role in the strong photothermoelectric photoresponse described in Chapter 2. Exploiting the efficiency of this energy flow process in graphene is a topic of current research [120].

4.3 Summary

As demonstrated above, the interaction mediated cascade of IE processes in doped graphene can lead to the generation of multiple electron-hole excitations by a single absorbed photon. A comparison with electron-phonon scattering indicates that the IE scattering is very efficient and can dominate carrier relaxation for typical dopings. Multiple pair excitations produced in the IE cascade triggered by a single photon feature an approximately linear scaling of the number of generated pairs and total cascade time with photo-excitation energy. These dependences, as well as a sharply peaked angular distribution of e-h pairs, provide clear experimental signatures for an IE-dominated cascade. Strong gate dependence of the cascade parameters affords a useful knob for the control of ultrafast scattering processes in graphene.

Crucially, the transfer of energy from photoexcited carriers to electronic degrees of freedom in graphene is efficient over a wide range of frequencies (from the UV to the infrared), unlike conventional semiconductor systems where the frequency range is limited by the band gap. Thus graphene enables enhanced quantum efficiencies and tunable energy transfer over a wide spectral range. This may enable a type of solar cell, called the "hot

carrier solar cell" [34] in which hot carriers can be directly extracted to provide efficiencies that beat the Shockley-Quiesser Limit [2].

Chapter 5

Shockley-Ramo Theorem & Long-Range Photocurrent in Gapless Materials

Many existing schemes of photodetection rely on transforming photon energy into electrical signals [121]. Photoresponse proceeds in three stages: 1) incoming radiation creates electron-hole pairs; 2) photoexcited pairs generate electric fields and charge movement in the system, inducing current in current-collecting contacts; 3) the induced current is amplified and converted to the output signal. Studies of photogalvanic effects typically focus on stage 1, which includes the phenomena occurring locally in the photoexcitation region (see e.g. Refs. [53, 122–125]). In contrast, stage 2 received relatively little attention. In this chapter[1] we discuss signal transduction in the system at stage 2, in particular the mechanisms of spatially non-local response.

As we will see, these mechanisms have much in common with the processes in charge detectors studied a long time ago by Shockley and Ramo in the context of vacuum-tube electronics. [35, 36, 126] They pointed out that the response of charge detectors is

[1]This chapter has been reproduced from JCW Song, LS Levitov, The Shockley-Ramo Theorem and Long-Range Photocurrent Response in Gapless Materials, *Physical Review B* [in production] Copyright (2014) by the American Physical Society.

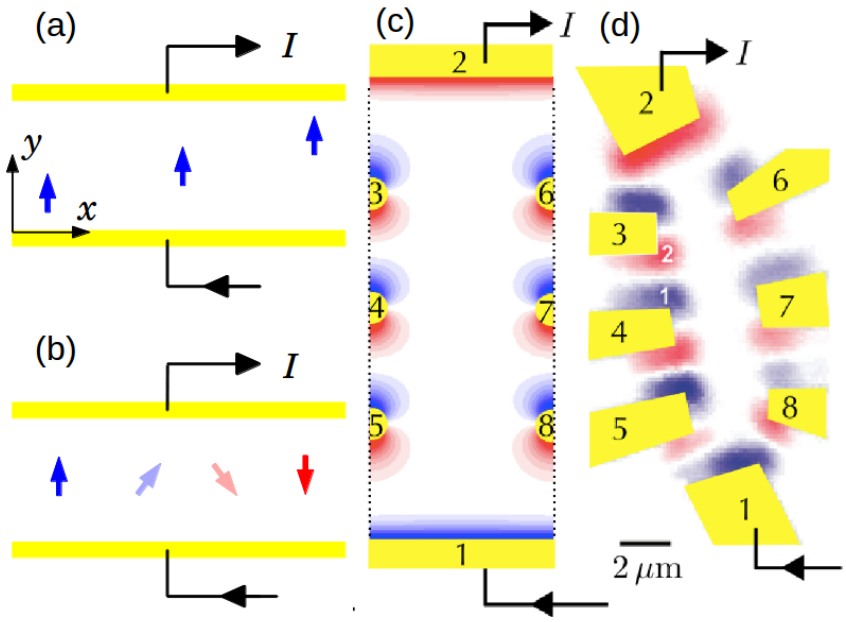

Figure 5.1: (a,b) Toy model for long-range photoresponse and directional effect in a strip $0 < y < w$ with current-collecting contacts at the sides $y = 0, w$ (see Sec 5.1). Different photocurrent sources $\mathbf{j}_{\rm ph}$ are schematically shown by arrows. The arrow color and intensity indicate the sign and magnitude of the induced net current I. The value I *does not depend* on the source position within the strip (a) but has strong dependence on its orientation (b). (c) Photocurrent pattern due to floating contacts that do not draw current (yellow semi-circles labeled 3-8). The photocurrent, drawn from contacts 1 and 2, is modeled as described in Sec. 5.3, see Eq.5.17,5.19.(d) Scanning photocurrent image of a 12 μm-long graphene device with six floating contacts 3-8. Note that the sign of photoresponse near floating contacts is correlated with the direction to the current-collecting contacts 1 and 2, but essentially independent of contact location within the system (data taken from Fig.2(a) of Ref. [37] with permission).

governed by long-range effects: The instantaneous electric currents induced by a moving charge are due to the electric field flux seen by each electrode rather than the amount of charge entering the electrode per second. As a result, the induced currents are only weakly sensitive to the charge position but depend strongly on the charge velocity magnitude and direction. The Shockley-Ramo (SR) approach—the seminal SR theorem—allows one to easily calculate the response. As we demonstrate, even though photoresponse in gapless materials originates from very different physics, it is described by a formalism similar to that of the SR theorem.

Spatial nonlocality of optoelectronic response is common for many gapped materials where it arises due to slow recombination of photoexcited carriers [121]. Recently, however, a long-range photocurrent response was reported in systems where carrier recombination is fast on the carrier diffusion timescales. Notably, this is the case in scanning photocurrent experiments that probe new gapless materials, such as graphene and topological insulators [5, 37, 38, 55, 67, 127, 128]. Photoresponse in these systems is of a distinctly global character: rather than being localized near current-collecting contacts, the photocurrent hot spots feature complex spatial patterns spanning the entire system area, typically separated by many microns from the contacts [37, 38, 55, 67, 127]. These large length scales may seem hard to reconcile with the short picosecond-scale recombination times over which the photoexcited carriers lose their energy and become part of the thermal distribution, traversing distances much less than system size.

The observed photoresponse also displays other striking features, in particular the *directional effect* (Fig.5.1). Namely, the photocurrent hot spots are highly sensitive to the orientation of inhomogeneities and interfaces, at which the hot spots are pinned, but totally independent of the distance from the contacts. The global character of photoresponse in combination with its strong dependence on the orientation relative to contacts is particularly striking in the data from Ref. [37] where this effect was first reported [reproduced in Fig.5.1(d)]. Here we introduce a framework that naturally explains how the non-locality can arise in the absence of slow recombination. This framework also provides a simple explanation for the directional effect.

5.1 The origin of the nonlocal and directional behavior

The reasons photoresponse in gapless materials is mediated by ambient carriers can be summed up as follows. On one hand, short recombination times lead to a rapid decay of the primary photoexcited carriers, preventing them from reaching contacts and directly contributing to photocurrent. On the other hand, ambient carriers can generate currents and fields reaching far from the photoexcitation spot. The main contribution to photoresponse is therefore an indirect one: a local photocurrent sets up an e.m.f. that drives ambient carriers outside the excitation region, and into the contacts.

These processes can be modelled by a spatially localized photogalvanic current $\mathbf{j}_{\rm ph}(\mathbf{r})$ induced by photoexcitation, and a diffusion current $\mathbf{j}_{\rm d}(\mathbf{r})$ due to ambient carriers in the material, obeying

$$\nabla \cdot (\mathbf{j}_{\rm d} + \mathbf{j}_{\rm ph}) = 0, \quad \mathbf{j}_{\rm d} = -\sigma(\mathbf{r})\nabla\phi, \tag{5.1}$$

where $\sigma(\mathbf{r})$ is position-dependent conductivity tensor, ϕ is the electrochemical potential. As we will see, the resulting response does not diminish with distance and displays the directional effect.

The origin of such a behavior can be understood by analyzing a special case: a spatially uniform system with constant conductivity. With regard to this toy model, some points of clarification are in order. First, on general symmetry grounds, local inhomogeneities, interfaces and boundaries are essential for generating photocurrent. Thus, a 'spatially uniform system' assumption only pertains to transport properties far outside the area where $\mathbf{j}_{\rm ph}$ is concentrated. Second, the assumption of spatial uniformity is used here merely to simplify the discussion. A more general situation will be analyzed in Sec. 5.2. Third, as we discuss in Sec. 5.3, photocurrent patterns are sensitive to the symmetries which govern photoresponse via a relation between $\mathbf{j}_{\rm ph}$ and local density gradients, see Eq.(5.19) and accompanying

discussion.

As a warm-up, we consider transport in an infinite 2D system in the presence a spatially localized photogalvanic current $\mathbf{j}_{\rm ph}(\mathbf{r})$. Fourier-transforming transport equations, Eq.(5.1), yields algebraic equations, giving a non-local relation

$$j_{{\rm d},i}(\mathbf{r}) = \int d^2r' D_{ik}(\mathbf{r},\mathbf{r}') j_{{\rm ph},k}(\mathbf{r}'), \tag{5.2}$$

$$D_{ik}(\mathbf{r},\mathbf{r}') = -\sum_q e^{i\mathbf{q}(\mathbf{r}-\mathbf{r}')} \frac{q_i q_k}{\mathbf{q}^2} = \frac{2n_i n_k - \delta_{ik}}{2\pi(\mathbf{r}-\mathbf{r}')^2}, \tag{5.3}$$

where $\mathbf{n}$ is a unit vector pointing from $\mathbf{r}'$ to $\mathbf{r}$.

Parenthetically, in writing Eq.(5.1) we make the usual assumptions that the magnetic effects are negligible and the electric field propagates instantaneously. Under these assumptions, the problem can be treated as electrostatics at each moment of charge movement (with the cutoff frequency value set by the retardation effects due to charge dynamics, see Eq.(5.16) below).

To link the power law found for D_{ik} and the global response we analyze a simple geometry: a strip $0 \le y \le w$ infinite in the x direction, with current-collecting contacts at the sides $y = 0, w$, as illustrated in Fig.5.1(a,b). We can extend the above analysis to explicitly evaluate the response induced by a localized source. As we will see, the net current flowing through the contacts equals

$$I = \frac{1}{w} \int d^2r' j_{{\rm ph},y}(\mathbf{r}'), \tag{5.4}$$

which depends on the y coordinate of $\mathbf{j}_{\rm ph}$ only. This result displays essential nonlocality since I is independent of $\mathbf{j}_{\rm ph}$ position [see Fig.5.1(a)]. While the independence of the x coordinate (in Eq. 5.4) follows directly from translational invariance, the independence of the y coordinate does not follow from any symmetry. It is counterintuitive and to a large degree comes as a surprise. Besides the 'global property' (independence of $\mathbf{j}_{\rm ph}$ position), our

result also displays the 'directional property' since the response depends on the y component of $\mathbf{j}_{\rm ph}$ only, reversing sign upon $\mathbf{j}_{\rm ph}$ reversal [see Fig.5.1(b)].

To derive Eq.(5.4), we note that the approach outlined in Eqs.(5.2),(5.3) can be reformulated in terms of the Greens function of Laplace's equation with zero boundary condition at $y = 0$, $y = w$,

$$D_{ik}(\mathbf{r}, \mathbf{r}') = -\nabla_i G(\mathbf{r}, \mathbf{r}')\nabla'_k, \quad \nabla^2 G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \tag{5.5}$$

where ∇ and ∇' are gradients with respect to $\mathbf{r}$ and $\mathbf{r}'$. Fourier-transforming with respect to x, we express the result through a 1D Greens function, $G(\mathbf{r}, \mathbf{r}') = \sum_q e^{iq(x-x')} g_q(y, y')$,

$$(\partial_y^2 - q^2) g_q(y, y') = \delta(y - y'). \tag{5.6}$$

Solving this equation in the interval $[0, w]$ with zero boundary conditions, we obtain

$$g_q(y, y') = A \sinh(q y_<) \sinh q(y_> - w), \tag{5.7}$$

where $y_< = \min(y, y')$, $y_> = \max(y, y')$, $A = \frac{1}{q \sinh(qw)}$. Plugging this in Eqs.(5.5),(5.2) and setting $y = 0$, we find normal current at the boundary, $j_n^{(\rm d)}(x) = j_{{\rm d},y}(x)_{y=0}$. We obtain

$$j_n^{(\rm d)}(x) = -\int d^2 r' \sum_q e^{iq(x-x')} \frac{\sinh q(y' - w)}{\sinh(qw)} \nabla' \cdot \mathbf{j}_{\rm ph}(\mathbf{r}'). \tag{5.8}$$

By mirror symmetry, only the component of $\mathbf{j}_{\rm ph}$ normal to the strip contributes to the above expression. Integration by parts gives

$$\begin{aligned} &\int_0^w dy' \frac{\sinh q(y' - w)}{\sinh(qw)} \partial_{y'} j_{{\rm ph},y}(y') = j_{{\rm ph},y}(y' = 0) \\ &- \int_0^w dy' \frac{q \cosh q(y' - w)}{\sinh(qw)} j_{{\rm ph},y}(y') \end{aligned} \tag{5.9}$$

The net current is evaluated as $I = \int dx (j_{{\rm d},y}(x) + j_{{\rm ph},y}(x))_{y=0}$. Using the relation $\int dx e^{iq(x-x')} = 2\pi\delta(q)$ we arrive at the result, Eq.(5.4), which displays the 'directional property' and the 'global property'.

It is instructive to note a relation between our calculation above and an electrostatic problem of a point dipole inserted in a parallel plate capacitor. The dipole induces image charges on the capacitor plates, which also display the directional property and the global property. Namely, the net induced charge values are given by $\Delta q_{1,2} = \pm\frac{1}{w}p\cos\theta$, where p and θ are the dipole magnitude and tilt angle, and w is the plate separation. The dependence of $\Delta q_{1,2}$ on θ and their independence of dipole position are identical to that for photoresponse, as illustrated in Fig.5.1(a,b). The origin of this relation can be traced to an isomorphism between the two problems, with $\mathbf{j}_{\rm d}$ and $\mathbf{j}_{\rm ph}$ playing the role of the electric field and dipole density in the electrostatic problem. As we will see in the next section, this result can be viewed as a special case of the SR theorem.

5.2 Mapping to the Shockley-Ramo problem

The global property and the directional property bear strong resemblance to the behavior in charge detectors described by the SR approach [35, 36, 126]. Before working out the connection between our problem and the SR approach, we briefly summarize Shockley and Ramo's results. Shockley and Ramo were concerned with the currents induced in the electrodes by charges moving in the free space inside a vacuum tube. The SR theorem provides a closed-form relation between the current induced by a moving charge e in the electrode k and the charge velocity and position, denoted by I_k, $\mathbf{v}(t)$ and $\mathbf{R}(t)$, respectively. The SR result, which is intrinsically nonlocal due to the long-range character of electric fields in vacuum, reads

$$I_k = e\mathbf{v}(t)\cdot\mathbf{E}_{\mathbf{r}=\mathbf{R}(t)}, \quad \mathbf{E}(\mathbf{r}) = \nabla w_k(\mathbf{r}), \tag{5.10}$$

The 'weighting potentials' $w_k(\mathbf{r})$ satisfy Laplace's equation with suitable boundary conditions on the electrodes ($w_k = 1$ at electrode k, and $w_k = 0$ at electrodes $j \neq k$). The SR

theorem is a foundation of ultra-fast charge sensing, such as particle detection in high energy physics [126, 129] and plasma diagnostics. [130] It can also be extended to charges moving in insulators [131].

In contrast, the relation between our problem and the SR-type treatment of charge detectors can be described as a mapping rather than merely an application of the SR approach to yet another system. This relation is based on an isomorphism between our problem and the SR problem, wherein the flow of ambient carriers and the photocurrent source play the role of electric field and moving charge, respectively. The long-range character of the response can be linked to charge continuity. The condition $\nabla \cdot \mathbf{j} = 0$ can be interpreted as incompressibility of current flow, with stream lines that do not terminate anywhere within the system. In addition, because the current is caused by a chemical potential gradient, the stream lines cannot form loops. This results in a response not diminishing with the distance between contacts and local photoexcitation, $\mathbf{j}_{\rm ph}$. As we show below, basically following the SR strategy, the system response can be described as

$$I = A \int \mathbf{j}_{\rm ph}(\mathbf{r}) \cdot \nabla\psi(\mathbf{r})\, d^2r, \tag{5.11}$$

where $\mathbf{j}_{\rm ph}(\mathbf{r})$ is local photogalvanic current in the photoexcitation region, ψ is a weighting field obtained by solving a suitable Laplace problem, and A is a prefactor which depends on device configuration (see Eq.(5.14)).

Spatial patterns predicted using Eq.(5.11) exhibit photocurrent-active structures with contrast which is essentially independent on their position within the system (see Figs.5.1,5.2,5.3). Such "global" photoresponse is known for one-dimensional systems, where Eq.(5.11) reduces to adding up the total potential drop across the device [18]. However, the generalized framework presented here yields photocurrent that can exhibit complex structures which are not anticipated in a one-dimensional approach.

We emphasize that the origin of nonlocality in our photoresponse problem is quite different from that in the SR problem, since the ambient carriers screen the long range electric field created by photoexcited carriers. As noted above, the nonlocality originates from long-range currents constrained by charge continuity relation. Further, the SR theorem is typically applied to high-speed charge detection, whereas we are concerned with the steady-state photocurrent. Yet, despite these differences, our approach yields a relation [Eq.(5.11)] which exhibits formal similarity with the SR theorem.

The starting point of our analysis is the continuity equation, Eq.(5.1). As discussed above, the two contributions to current in Eq.(5.1) have very different spatial dependence: the photogalvanic current $\mathbf{j}_{\rm ph}$ is present in the excitation region, whereas the diffusion current $\mathbf{j}_{\rm d}$ is nonzero throughout the entire material. Below we focus on the simplest situation when transport can be described by a position-dependent 2×2 conductivity tensor $\sigma(\mathbf{r})$. The diffusion current satisfies the usual relation $\mathbf{j}_{\rm d} = -\sigma(\mathbf{r})\nabla\phi$, where ϕ is the electrochemical potential. The boundary conditions in this transport problem are zero current through the sample boundary, $\mathbf{n} \cdot (\mathbf{j}_{\rm d} + \mathbf{j}_{\rm ph}) = 0$, and constant potential at the contacts, $\mathbf{n} \times \nabla\phi = 0$ (here $\mathbf{n}$ is the normal to the boundary).

To handle the non-local response, we introduce an auxiliary weighting field $\psi(\mathbf{r})$ in the bulk of the material, satisfying

$$\nabla \cdot \mathbf{j}^{(\psi)}(\mathbf{r}) = 0, \quad \mathbf{j}^{(\psi)} = -\sigma^{\rm T}\nabla\psi, \tag{5.12}$$

where $\sigma^{\rm T}$ is a 2×2 matrix transposed to σ. The field $\psi(\mathbf{r})$ satisfies appropriate boundary conditions at the boundary and contacts, $\mathbf{n} \cdot \mathbf{j}^{(\psi)}(\mathbf{r}) = 0$ and $\mathbf{n} \times \nabla\psi(\mathbf{r}) = 0$, respectively (here $\mathbf{n}$ is a normal unit vector at the boundary). Multiplying the continuity equation for the physical current $\mathbf{j}_{\rm d} + \mathbf{j}_{\rm ph}$ by $\psi(\mathbf{r})$, integrating over the sample area, and using Gauss'

theorem, we obtain

$$\int \nabla\psi(\mathbf{r})\cdot\mathbf{j}_{\rm ph}(\mathbf{r})d^2r = \sum_k \psi_k I_k - \phi_k I_k^{(\psi)} \tag{5.13}$$

where k labels contacts. The quantities on the right hand side are the net currents flowing in each of the contacts, $I_k = \int_{C_k} \mathbf{n}\cdot\mathbf{j}_k d\ell$, and potentials on these contacts. We emphasize that Eq.(5.13) holds on very general grounds regardless of whether a particular contact is drawing current ($I_k \neq 0$) or is floating ($I_k = 0$). The expression on the left hand side depends on the microscopic distribution $\mathbf{j}_{\rm ph}(\mathbf{r})$ inside the material, whereas the expression on the right hand side is a function of currents and potentials at the contacts, thereby providing a general relation between position-dependent photoexcitation and the measured photocurrent.

It is convenient to choose $\psi(\mathbf{r})$ such that $I_k^{(\psi)} = 0$ for all floating contacts. Then the contribution to Eq.(5.13) due to floating contacts drops out entirely, yielding a relation which only includes the contacts that actually draw current. It is also straightforward to account for the effect of an external circuit. We consider the current drawn through a pair of contacts 1 and 2 (see Fig.5.1) and write $I_{1(2)}^{(\psi)} = \mp(\psi_1-\psi_2)/R$, $I_{1(2)} = \pm(\phi_1-\phi_2)/R_{\rm ext}$, with R and $R_{\rm ext}$ the resistance of the sample and of the external circuit, respectively. Setting $\psi_1 - \psi_2 = 1$, we obtain Eq.(5.11) with the prefactor

$$A = R/(R + R_{\rm ext}). \tag{5.14}$$

Despite its apparent simplicity, Eq.(5.11) accounts for all the key effects that impact photoresponse, such as system geometry, structure, inhomogeneity, etc. Similar to the canonical SR relation, Eq.(5.10), the relation in Eq.(5.11) is essentially nonlocal due to the long-range character of currents in the system.

Here we briefly discuss the validity of our approach. Our transport equations, Eq.(5.1), are written in a quasistatic approximation. This is similar to the SR approach which treats the electric field induced by a moving charge as instantaneous. The SR result

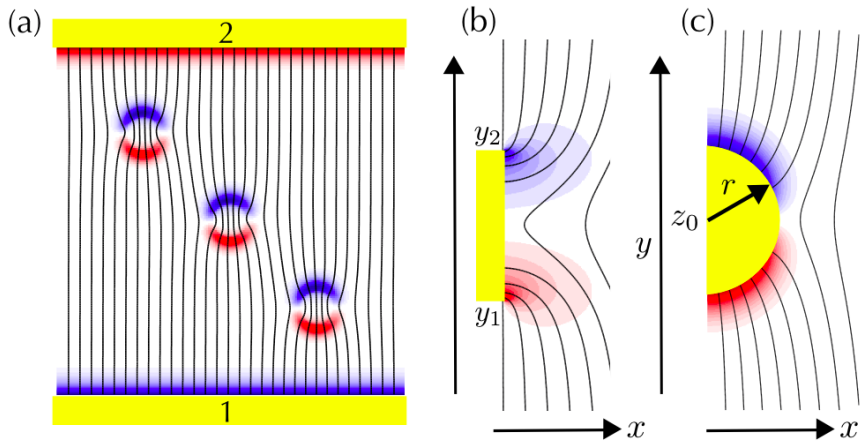

Figure 5.2: Directional effect in photoresponse accounting fully for the distortions of the weighting field. (a) Photocurrent pattern due to three circular regions, modeled in the same way as in Fig.5.3 (b). The conductivity inside each region is taken to be 10 times larger than the background conductivity. (b,c) Photoresponse and the field lines for $\nabla\psi$ near floating contacts of two different shapes, a rectangle and a semicircle, obtained using the conformal mapping approach, Eq.(5.17).

is therefore valid at frequencies below the cutoff set by the EM retardation timescale, $\omega \ll \omega_0 = c/L$, where L is system size. In our case, the cutoff frequency is set by the characteristic time for charge dynamics in the system. An estimate below yields very short timescales, i.e. a very fast response.

A crude estimate of timescales can be obtained by reinstating the time dependent term in the continuity equation. For a spatially uniform system, the dynamics of the Fourier harmonics of charge density is given by

$$\partial_t \delta n_{\mathbf{k}}(t) = -\frac{2\pi}{\kappa}\sigma|\mathbf{k}|\delta n_{\mathbf{k}}(t), \tag{5.15}$$

where σ is the sheet conductivity per square area and κ is the dielectric constant. For a simple estimate, taking parameter values $|\mathbf{k}| \approx \pi/L$, $L = 10\,\mu\text{m}$, $\kappa = 5$, $1/\sigma = 1\,\text{k}\Omega$, we obtain a sub-picosecond response time

$$\tau = \kappa L/(2\pi^2\sigma) \approx 0.3\,\text{ps}, \tag{5.16}$$

which is considerably shorter than typical cooling and recombination times in graphene. Fast response makes the photocurrent a potentially useful probe for the dynamical processes in the excitation region. It also makes gapless materials viable for applications in high-speed optoelectronics.

5.3 Geometry of the weighting field

The general features of Eq.(5.11) can be illustrated for a spatially uniform system of a rectangular shape. In this case, the weighting field $\psi(\mathbf{r})$ is a linear function, $\nabla\psi = \hat{\mathbf{y}}/L$, with L the system length. Constant $\nabla\psi$ yields Eq.(5.4) derived in Sec. 5.1 by a direct calculation. As discussed above, this describes a response which is invariant upon spatial translation of $\mathbf{j}_{\rm ph}(\mathbf{r})$ (the global property). At the same time, the sign and the magnitude of the response depend on the angle between $\nabla\psi(\mathbf{r})$ and $\mathbf{j}_{\rm ph}(\mathbf{r})$ (the directional effect).

In order to test the robustness of this behavior, we now proceed to analyze a more realistic situation where spatial inhomogeneity in conductivity $\sigma(\mathbf{r})$ is essential. To analyze the inhomogeneous problem, we can use a numerical procedure to obtain the exact profile $\psi(\mathbf{r})$. Fig.5.2 (a) shows photocurrent patterns from three circular regions with a mismatch between the inner and outer conductivity, which causes significant distortions of the $\nabla\psi$ field lines. Yet these distortions do not impact the overall trends discussed above, the global character of the response and the directional effect.

In contrast, the weighting field distortions have a very dramatic effect near contacts. Even if a contact does not draw net current, it *short-circuits* the current flowing in its vicinity, leading to a non-vanishing normal component of $\nabla\psi$ near the surface of a contact (see Fig. 5.2). For $\mathbf{j}_{\rm ph}$ which is normal to the contact, this gives a nonzero, sign-changing photoresponse, as in Fig.5.1 (c,d).

For ideal contacts, the field ψ can be found using the conformal mapping approach, giving $\psi(\mathbf{r}) = A \operatorname{Im} w(z)$. Here w is a suitable analytic function of a complex variable $z = x + iy$, which satisfies the equipotential condition at the contact surface. We illustrate this for a flat contact and for a semicircular contact (see Fig.5.2 (b,c)):

$$w_{\mathrm{b}}(z) = \sqrt{(z - y_1)(z - y_2)}, \quad w_{\mathrm{c}}(z) = \tilde{z} - r^2/\tilde{z}, \tag{5.17}$$

$\tilde{z} = z - z_0$, where the flat contact is positioned at $y_1 < y < y_2$, $x = 0$, and the semicircular contact is of radius r and is positioned at $z = z_0$. We assume that the contacts are floating and are small compared to the system size. At large z, ψ asymptotically approaches the linear dependence $\psi \propto y$ found above. The photocurrent at the contact is proportional to $\mathbf{n} \cdot \nabla\psi$. For the flat contact,

$$\partial_x \psi(\mathbf{r})_{x=0} = A \frac{y - \frac{1}{2}(y_1 + y_2)}{\sqrt{(y - y_1)(y_2 - y)}}, \quad y_1 < y < y_2. \tag{5.18}$$

Since this quantity is an odd function of $y - \frac{1}{2}(y_1 + y_2)$, the net current drawn in the contact vanishes, as appropriate for a floating contact. Similar sign-changing behavior is found for the semicircular contact, see Fig.5.2 (c). The sign-changing pattern is oriented in such a way that the parts showing high photoresponse are facing the contacts through which the photocurrent is drawn. This behavior is in agreement with the directional effect, see Fig.5.1(d).

Next, we discuss application of our approach for diagnostic of different types of photogalvanic response. The value $\mathbf{j}_{\mathrm{ph}}(\mathbf{r})$ depends on system properties in the photoexcitation region. By symmetry, no photogalvanic effect can occur in a spatially uniform system (assuming unpolarized light). In the presence of a density gradient $\nabla n(\mathbf{r})$, the local photogalvanic current can be described as

$$\mathbf{j}_{\mathrm{ph}}(\mathbf{r}) = [\alpha \nabla n(\mathbf{r}) + \beta \hat{\mathbf{z}} \times \nabla n(\mathbf{r})]\, J(\mathbf{r}), \tag{5.19}$$

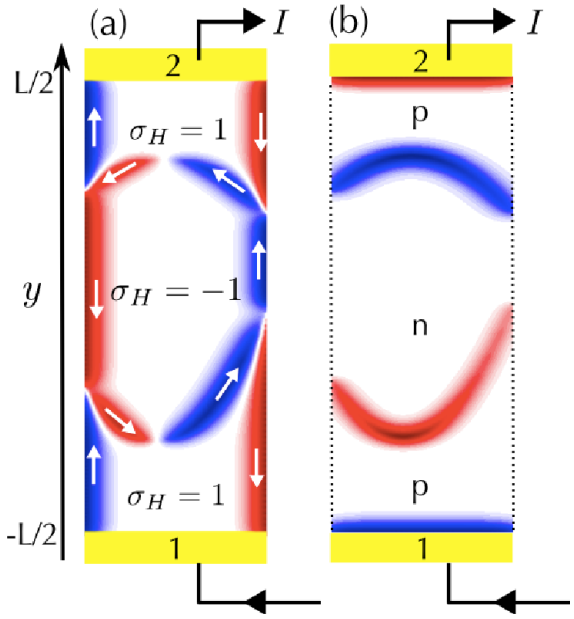

Figure 5.3: Scanning photocurrent images (from our model) for different mechanisms of photoresponse. The photocurrent, drawn from contacts 1 and 2, is modeled by Eqs.(5.11),(5.19). (a) Photocurrent pattern in a chiral material, where $\sigma_H = \pm 1$ marks regions of different chirality. Local photocurrent direction is governed by edge states (white arrows). (b) Photocurrent pattern in a non-chiral system with a step-like density inhomogeneity (see text).

where α and β are material constants, and $J(\mathbf{r})$ is the absorbed optical power. In general, α is finite in all materials, whereas β is only non-zero in *chiral systems* where edge-state transport allows $\mathbf{j}_{\rm ph}$ to be directed along the contours of $n(\mathbf{r})$. This is the case in chiral materials such as topological insulators due to coupling between orbital motion and spin [123, 124, 128], or in non-chiral materials in the presence of a magnetic field [55].

The effects of spatial inhomogeneity are illustrated in Fig.5.3 for the chiral response (a) and the nonchiral response (b). The patterns in Fig.5.3 (a),(b) were obtained using a spatially uniform weighting field approximation, $\nabla\psi \approx \hat{\mathbf{y}}/L$. For illustrative purposes, we use a step-like density profile, with n taking one value in the middle region and another value in the top and bottom regions, identical for (a) and (b). For (a) we use $\mathbf{j}_{\rm ph}$ with $\alpha = 0$ and finite β, for (b) it is the other way around. In both cases, the photocurrent is zero in the regions of constant n, and nonzero near the steps. The differences in the sign and

magnitude of the response reflect the fundamental difference in physics in the cases (a) and (b).

Model (a) describes photoresponse in chiral systems arising at the interfaces between domains of opposite chirality. Physically, it may represent a quantum Hall system near a plateau transition [132], or a system in which nonzero chirality results from spontaneous ordering [133]. The different signs of chirality, labelled by $\sigma_H = \pm 1$ in Fig.5.3(a), can be associated with the clockwise and counter-clockwise edge states, labelled by white arrows. Notably, the sign and magnitude of photocurrent depend on the direction of current flow in the edge states. The photocurrent is also nonzero at system boundaries, indicating the presence of current carrying edge states. This can be used to identify the edge states and domains with different chirality in experiment.

Fig.5.3 (b) shows the non-chiral photocurrent response for the same density profile as in Fig.5.3 (a). Physically, (b) may describe systems such as graphene with spatial inhomogeneity giving rise to p-n boundaries separating regions with electron-like and hole-like polarity [67]. In this case, $\mathbf{j}_{\rm ph}$ is normal to the contours of $n(\mathbf{r})$, making the sign and magnitude of the response dependent on the orientation of the interfaces viz. $\hat{\mathbf{y}} \cdot \mathbf{j}_{\rm ph}$. Also, since $\mathbf{j}_{\rm ph}$ is normal to boundaries whereas $\nabla\psi$ is tangential, the photocurrent vanishes at the system edge.

A very different behavior is found near contacts, since $\nabla\psi$ is normal to the contact surface, see Fig.5.1 (c). In this case, a nonzero response arises both near the contacts through which current is drawn and near floating contacts (see also Fig.5.2). Notably, the response depends on the floating contact orientation but not on its position within the system. This is in agreement with experimental observations of Ref. [37], which are reproduced in Fig. 5.1 (d). All photocurrent patterns in Fig.5.1 and Fig.5.3, despite their different physical origin, share two common trends: strong directional sensitivity and global character (posi-

tional independence). This behavior makes the photocurrent patterns particularly useful in identifying symmetry breaking and inhomogeneity in gapless materials.

5.4 Conclusions

In summary, our approach explains several puzzling aspects of photocurrent response in gapless materials, in particular the striking non-locality and the directional effect observed in Ref. [37]. By analyzing different mechanisms of photoresponse, we demonstrate that it is uniquely capable of revealing spatial patterns arising due to symmetry breaking, chirality, or inhomogeneities. Fast photoresponse makes gapless materials potentially useful for a variety of high-speed electronics applications. Sub-picosecond response times estimated for graphene make photoresponse a useful probe of carrier dynamics in this material.

Chapter 6

Energy-driven Drag in Graphene

Vertical heterostructures comprised of a few graphene layers separated by an atomically thin insulating layer [134] afford new ways to probe the effects of electron interactions at the nanoscale. Typical layer separation d in these structures (1-2 nm) can be very small compared to the characteristic electron lengthscales such as the de Broglie wavelength, λ, and the screening length. This defines a new strong-coupling regime, $d \ll \lambda$, wherein the interlayer and intralayer interactions are almost equally strong. Fast momentum transfer between electron subsystems in the two layers and strong Coulomb drag have been predicted in this regime [135–140] with characteristic dependence on doping, temperature and layer separation distinct from that in previously studied systems [46].

Recent measurements [44, 141], while confirming theoretical predictions away from charge neutrality (CN), yield unexpected results at CN. Conventional momentum drag (P-mechanism) vanishes at CN because the sign of P-mechanism depends on the polarity of charge carriers [142]. However, experiment [44] shows a *sharp peak* in the drag response at CN. This disparity indicates that new physics is involved in driving drag in graphene at CN.

In this chapter[1], we propose a new mechanism for drag: energy-driven drag (E-

[1]Reproduced in part from JCW Song, LS Levitov, Energy-Driven Drag at Charge Neutrality in Graphene, *Physical Review Letters* **109**, 236602 Copyright (2012) by the American Physical Society.

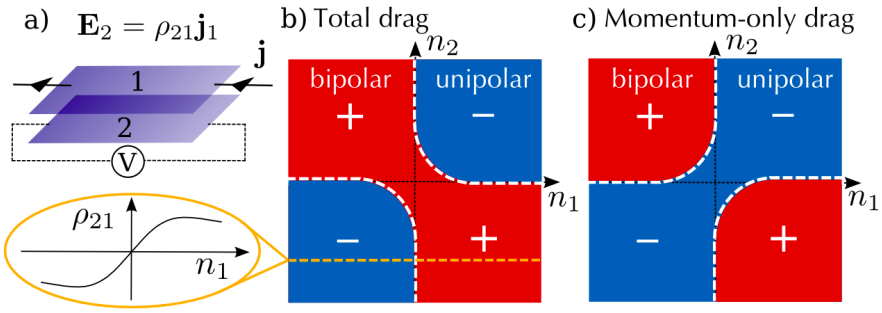

Figure 6.1: Different mechanisms for Coulomb drag in graphene heterostructures. E-mechanism dominates over P-mechanism near zero doping, whereas P-mechanism dominates at higher doping. The sign of the drag response depends on carrier polarity (a). For potential fluctuations of equal sign in the two layers, Eq.(6.1), the net drag (b) features a pair of nodal lines (white dashed lines). Positive drag in the avoided crossing region at zero doping is dominated by E-mechanism. The resulting dependence is distinct from P-mechanism-only drag (c) smeared by correlated density fluctuations, $\delta\mu_1 \approx \delta\mu_2$.

mechanism). As we will show, E-mechanism becomes important in the *adiabatic regime* where the electronic system is thermally decoupled from the lattice (hot carrier regime). In graphene, slow electron-lattice cooling means that thermal decoupling persists over few-micron length scales even at room temperature [11,12,26]. As a result, electronic heat current is a dynamical variable, that together with charge current, governs transport behavior.

When $d \ll \lambda$, the interlayer electron-electron scattering mediates efficient vertical energy transfer between layers, which couples electron temperatures in the layers. In the adiabatic regime, coupled lateral energy flow in the two electronic systems, via thermoelectric effect, yields nonzero drag (see below). E-mechanism predicts drag which has a characteristic density dependence (illustrated in Fig. 6.1(b)) featuring a positive drag resistivity at double neutrality, $\rho_{21} > 0$. The sign, as well as the peak structure in ρ_{21}, agrees with experiment [44].

E-mechanism arises due to the coupling between vertical energy transfer and lateral charge and energy transport via spatial density inhomogeneity which is intrinsic to graphene. Density inhomogeneity is known to be particularly strong at CN in the electron-

hole puddle regime [70], providing the dominant disorder potential in clean samples. When a charge current is applied in layer 1, density inhomogeneity produces spatially varying heating/cooling [see Eq.(6.2)]. Strong thermal coupling between the electron systems in the two layers, mediated by the interlayer energy transfer, leads to a temperature pattern in layer 2 that tracks that in layer 1, $\delta T_2(r) \approx \delta T_1(r)$. Further, since the disorder correlation length $\xi_{\rm dis}$ can reach 100 nm in G/BN heterostructures [143, 144], exceeding the layer separation by orders of magnitude, the potential fluctuations are nearly identical in the two layers,

$$\langle \delta\mu_1(r)\delta\mu_2(r')\rangle > 0 \tag{6.1}$$

for $r \approx r'$. As a result, the position-dependent thermopower induced by the gradient $\nabla\delta T_2(r)$ is correlated with the heating/cooling pattern in layer 1, giving rise to a nonzero ensemble-averaged drag voltage in layer 2.

Our mechanism predicts a particular sign of the energy contribution to drag. As a result, the density dependence for the net drag (E- and P-mechanism combined) features a split-up pattern of nodal lines with an "avoided crossing" at zero doping, as illustrated in Fig. 6.1 (b). The double sign change along the main diagonal $n_1 = n_2$ and the peak at $n_{1,2} = 0$ make E-mechanism easy to distinguish experimentally.

As a parenthetical remark, the correlated density inhomogeneity, Eq.(6.1), also affects the P-mechanism, however its effect is opposite to that of the E-mechanism. If P-mechanism were the dominant contribution near zero doping, the pattern of nodal lines would be such that the drag sign was constant along the main diagonal (see Fig.6.1b and c).

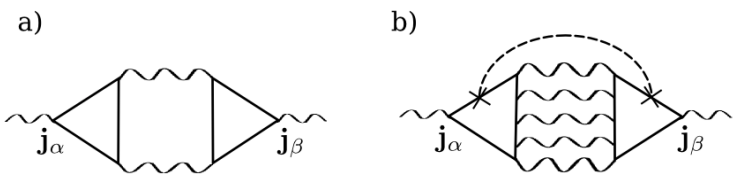

Figure 6.2: Feynman diagrams for P-mechanism (a) and E-mechanism (b) for drag. Wavy lines represent interactions, dashed line represents disorder averaging. The ladder in (b) represents a long-wavelength charge-neutral mode.

6.1 Energy-driven mechanism for Drag

E-mechanism is mediated by neutral modes (particle-hole excitations, or electron-lattice temperature imbalance) which are of a long-range character, and thus can be described by a hydrodynamic approach. The relevant length scales for these modes are ξ_T and $\xi_{\rm dis}$, the electron-lattice cooling length and the inhomogeneity correlation length, respectively. For a long-range disorder potential and not too low temperatures, the lengthscales ξ_T and $\xi_{\rm dis}$ are larger than the inelastic mean free path, $\ell = v/\gamma$, where γ is the electron-electron scattering rate. As a result, E-mechanism is captured by a hydrodynamic framework which involves charge current $\mathbf{j}$ and heat current $\mathbf{j}_{\rm q}$, which in the ballistic transport regime are related by

$$\mathbf{j}_{\rm q}(\mathbf{r}) = Q(n)\mathbf{j}, \quad Q[n(\mathbf{r})] = \mathcal{S}[n(\mathbf{r})]T/e, \tag{6.2}$$

where $\mathcal{S}(n)$ is the entropy per carrier, $n(\mathbf{r})$ is the density profile, and $e < 0$ is the carrier charge. In the ballistic regime, using the electron temperature approximation, we find (see Sec. 6.4)

$$Q = \frac{2\pi^2 k_B^2 T^2 \mu}{3e(\mu^2 + \Delta^2(T))}. \tag{6.3}$$

where $\Delta(T)$ accounts for the Dirac point broadening due to disorder and thermal fluctuations.

It is instructive to compare the Feynman diagrams describing different mechanisms

(see Fig.6.2). The characteristic momenta are fairly large for P-mechanism ($\sim k_F$), making it a local contribution. In contrast, E-mechanism includes ladder diagrams representing long-range modes propagating over distances of order $\xi_T \gg k_F^{-1}$.[2]

To illustrate the relation between energy and charge transport, we first analyze in-plane resistivity in a single layer. According to Eq.(6.2), spatial inhomogeneity leads to heating/cooling in the presence of uniform charge current (as in the Joule-Thomson process). The spatial temperature profile can be found from $-\nabla\kappa\nabla\delta T + \lambda\delta T = -\nabla\cdot\mathbf{j}_{\rm q}$, where κ is the thermal conductivity and $\lambda\delta T$ is the electron-lattice cooling power. A temperature gradient $\nabla\delta T$ drives thermopower, providing additional dissipation and thereby increasing resistivity. Onsager reciprocity combined with Eq.(6.2) gives $\mathbf{E}(r) = (Q[n(\mathbf{r})]/T)\nabla\delta T$ (see section below). Taking an ensemble average over small density fluctuations, $\delta\mu \ll k_B T, \mu$, we find an increase in the in-plane resistivity, $\langle\rho_{\alpha\beta}\rangle = \rho^0_{\alpha\beta} + \Delta\rho_{\alpha\beta}$, $[\alpha(\beta) = x, y]$, where

$$\Delta\rho_{\alpha\beta} = \frac{1}{T}\sum_{|\mathbf{q}|\lesssim 1/\ell} \frac{\langle\delta Q(-\mathbf{q})\delta Q(\mathbf{q})\rangle}{\kappa q^2 + \lambda} q_\alpha q_\beta. \tag{6.4}$$

Since the derivative $\partial Q/\partial\mu$ peaks at $\mu = 0$, this results in $\Delta\rho_{\alpha\beta}$ that peaks at CN. The temperature dependence estimated below is $\Delta\rho \propto T^2$, reminiscent of super-linear power laws for resistivity frequently observed at small doping [145]. A contribution of nonthermal modes to $\Delta\rho$ was analyzed in Ref. [146].

Generalizing this analysis to two layers coupled by vertical energy transfer and accounting for correlated density fluctuations, Eq.(6.1), we find an ensemble-averaged drag response $\mathbf{E}_2 = \rho_{21}\mathbf{j}_1$,

$$\rho_{21}^{(\rm e)} = \frac{1}{2T\tilde{\kappa}}\frac{\partial Q}{\partial\mu_1}\frac{\partial Q}{\partial\mu_2}\sum_{\mathbf{q}}\frac{\langle\delta\mu_2(-\mathbf{q})\delta\mu_1(\mathbf{q})\rangle}{1+\tilde{\ell}^2\mathbf{q}^2}. \tag{6.5}$$

Here $\tilde{\kappa} = \kappa_1 + \kappa_2$ is the net thermal conductivity of the two layers, μ is the chemical potential,

[2]We note that we will not be calculating these diagrams explicitly. Instead, we use a hydrodynamical approach which makes the physical origin of the E-mechanism plain.

and $\tilde{\ell}$ is the interlayer cooling length. This length is estimated below and is shown to be of order of the inelastic mean free path, $\tilde{\ell} \sim \ell$, much shorter than the electron-lattice cooling length ξ_T. Because the sign of the correlator in Eq.(6.5) is positive, energy-driven drag has the same sign as $\Delta\rho_{\alpha\beta}$ in Eq.(6.4), i.e. is *positive at zero doping*. This results in a double sign change along the main diagonal $n_1 = n_2$, as pictured in Fig. 6.1 (b). The density dependence for $\rho_{21}^{(\mathrm{e})}$ features a peak at zero doping (see Fig.6.3) which is a hallmark of the E-mechanism regime.

Positive correlation, Eq.(6.1), is expected for disorder potential dominated by charge impurities [118, 119, 147]. For the correlator $\langle \delta\mu_1 \delta\mu_2 \rangle$ of a negative sign, conjectured for strain-induced charge puddles [148], our analysis predicts a negative drag at zero doping. Hence drag is a useful tool for probing the origin of inhomogeneity in graphene.

6.2 Interlayer Energy Relaxation

We begin by studying the energy transfer mediated by the Coulomb interaction between the electronic systems in the two layers (Fig.6.1(a)). This is described by the Hamiltonian

$$\mathcal{H} = \sum_i \int d^2\mathbf{r} \psi_i^\dagger(\mathbf{r}) \Big[-i\hbar v \sigma \cdot \nabla + \delta\mu_i(\mathbf{r}) \Big] \psi_i(\mathbf{r}) + \mathcal{H}_{\mathrm{el-el}} \tag{6.6}$$

where $i, j = 1...2N$ index layer, and spin/valley degrees of freedom, $\delta\mu(\mathbf{r})$ describes the slowly varying disorder potential, v is the Fermi velocity. The electron-electron interactions are described by $\mathcal{H}_{\mathrm{el-el}} = \frac{1}{2} \sum_{\mathbf{q},\mathbf{k},\mathbf{k}',i,j} V_{ij}(\mathbf{q}) \psi^\dagger_{\mathbf{k}+\mathbf{q},i} \psi^\dagger_{\mathbf{k}'-\mathbf{q},j} \psi_{\mathbf{k}',j} \psi_{\mathbf{k},i}$.

In our analysis, we ignore the correction due to finite layer separation d, approximating the interlayer interaction by the bare Coulomb interaction, $V_{ij}(\mathbf{q}) \approx V^0_{\mathbf{q}} = 2\pi e^2/\varepsilon|\mathbf{q}|$ with ε the background dielectric constant. This approximation is valid when the lengthscale d is small compared to the screening length and Fermi wavelength in the layers, which is the case for systems of interest [134]. The random-phase approximation then yields a screened

interaction $V_{ij}(\mathbf{q}) = V^0_{\mathbf{q}}/[1 - V^0_{\mathbf{q}}(\Pi_1(\mathbf{q},\omega) + \Pi_2(\mathbf{q},\omega))]$ for i, j in different layers.

6.2.1 Calculating the energy relaxation rate

We describe the energy distribution of carriers in each layer by a Fermi distribution at temperatures $T_{1,2}$. Using Fermi's golden rule we can calculate the rate of energy exchange between the two layers.

$$W_{\mathbf{k_1}',\mathbf{k_1}} = \frac{2\pi N}{\hbar}\sum_{\substack{\mathbf{k_2},\mathbf{k_2}'\\ \mathbf{q}}} F_{\mathbf{k_2},\mathbf{k_2}'}|V_{\mathbf{q}}|^2 f(\epsilon_{\mathbf{k_2}})[1 - f(\epsilon_{\mathbf{k_2}'})]\delta_\epsilon\delta_1\delta_2 \tag{6.7}$$

where $\delta_\epsilon = \delta\big(\epsilon_{\mathbf{k_1}',\mathbf{k_1}} + \epsilon_{\mathbf{k_2}',\mathbf{k_2}}\big)$, and $\delta_1 = \delta_{\mathbf{k_1}',\mathbf{k_1}+\mathbf{q}}$ and $\delta_2 = \delta_{\mathbf{k_2}',\mathbf{k_2}-\mathbf{q}}$. Here $N = 4$ is the number of spin/valley flavors, $\{1, 2\}$ denote the different layers, $\epsilon_{\mathbf{k}',\mathbf{k}} = \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}$, and $F_{\mathbf{k},\mathbf{k}'} = |\langle\mathbf{k}'\alpha|\mathbf{k}\beta\rangle|^2$ is the coherence factor (α, β label states in the electron and hole Dirac cones). $V_{\mathbf{q}}$ is the screened inter-layer Coulomb interaction described below. The energy-loss power is

$$\mathcal{J} = N\sum_{\mathbf{k_1},\mathbf{k}_1'} W_{\mathbf{k_1}',\mathbf{k_1}}(\epsilon_{\mathbf{k_1}} - \epsilon_{\mathbf{k}_1'})f(\epsilon_{\mathbf{k_1}})\big[1 - f(\epsilon_{\mathbf{k}_1'})\big]F_{\mathbf{k_1},\mathbf{k_1}'} \tag{6.8}$$

We can simplify the evaluation of these sums by writing $\delta\big(\epsilon_{\mathbf{k_1}',\mathbf{k_1}}+\epsilon_{\mathbf{k_2}',\mathbf{k_2}}\big) = \int_{-\infty}^{\infty} d\omega\delta\big(\epsilon_{\mathbf{k_1}',\mathbf{k_1}} - \omega\big)\delta\big(\omega+\epsilon_{\mathbf{k_2}',\mathbf{k_2}}\big)$ and using the identity $f_s(\epsilon_{\mathbf{k}})\big[1-f_s(\epsilon_{\mathbf{k}'})\big] = \big(f_s(\epsilon_{\mathbf{k}})-f_s(\epsilon_{\mathbf{k}'})\big)\times\big(N_s(\epsilon_{\mathbf{k}',\mathbf{k}})+1\big)$ where $N(\omega) = 1/(e^{\omega/k_BT} - 1)$ is the Bose function taken at the electron temperature (of that particular layer) and $s = \{1, 2\}$ denotes the layers. Using the quantities

$$\chi_s''(\mathbf{q},\omega) = N\sum_{\mathbf{k}} F_{\mathbf{k},\mathbf{k}+\mathbf{q}}\big(f_s(\epsilon_{\mathbf{k}}) - f_s(\epsilon_{\mathbf{k}+\mathbf{q}})\big)\delta\big(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \omega\big). \tag{6.9}$$

where $\chi_s''(\mathbf{q},\omega) = -\frac{1}{\pi}\mathrm{Im}\Pi(\mathbf{q},\omega)$ is the imaginary part of the susceptibility and $\Pi(\mathbf{q},\omega) = N\sum_{\mathbf{k}} F_{\mathbf{k},\mathbf{k}+\mathbf{q}}\frac{f_s(\epsilon_{\mathbf{k}})-f_s(\epsilon_{\mathbf{k}+\mathbf{q}})}{\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}+\mathbf{q}}+\omega+i0}$ is the polarization operator. Using these we can re-write the

energy loss power as[3]

$$\mathcal{J} = \frac{\pi}{\hbar}\int_{-\infty}^{\infty} d\omega\omega\big(N_2(\omega) - N_1(\omega)\big)\sum_{\mathbf{q}} |V_{\mathbf{q}}|^2 \chi_1''(\mathbf{q},\omega)\chi_2''(\mathbf{q},\omega) \tag{6.10}$$

where we have noted that $N(-\omega) = -(1 + N(\omega))$ and used the fact that $\chi_s''(\mathbf{q},\omega) = -\chi_s''(-\mathbf{q},-\omega)$ so as to only keep the odd part of the product $N_2(-\omega)(N_1(\omega)+1)$.

The energy transfer between the two layers is dependent on the Coulomb interaction $V_{\mathbf{q}}$ between the layers. We treat $V_{\mathbf{q}}$ by accounting for polarization in both layers and screening in the RPA approximation. The RPA-screened coulomb interaction is

$$V_{\mathbf{q}} = \frac{V_{\mathbf{q}}^0}{1 - V_{\mathbf{q}}^0(\Pi_1(\mathbf{q},\omega) + \Pi_2(\mathbf{q},\omega))}, \quad V_{\mathbf{q}}^0 = \frac{2\pi e^2}{\varepsilon q} \tag{6.11}$$

where ε is the background dielectric constant. Here we ignored the correction due to finite interlayer spacing d, approximating the interlayer interaction $V_{\mathbf{q}} = \frac{2\pi e^2}{\varepsilon q}e^{-d|q|} \approx V_{\mathbf{q}}^0$. This approximation is valid when the layer separation d is small compared to the screening length in the layers, $dq_0 \ll 1$, $q_0 = -\Pi_1(\mathbf{q}=0) - \Pi_2(\mathbf{q}=0)$. We will be interested in the regime when d is small compared to the Fermi wavelength in the layers, for which the above approximation is adequate.

In the degenerate limit, $\mu \gg k_B T$, the polarization is given by

$$\Pi(q,\omega) = -\nu(\mu)\left(1 - \frac{\omega}{\sqrt{(\omega+i0)^2 - q^2v^2}}\right) \tag{6.12}$$

In the limit of $q \ll q_0$ where $q_0 = (2\pi e^2/\varepsilon)(\nu_1 + \nu_2)$, where $\nu_{1,2}$ is the density of states at the Fermi level in each layer, we can write $|V_q| = 1/|\Pi_1(\mathbf{q},\omega) + \Pi_2(\mathbf{q},\omega)|$. Noting that in the degenerate limit, intra-band transitions are the dominant processes we approximate $\chi''(\mathbf{q},\omega)$ by the imaginary part of Eq.(6.12). This allows us to write the cooling power as

$$\mathcal{J}_{12} = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} d\omega\omega\big(N_2(\omega) - N_1(\omega)\big)\sum_{|\mathbf{q}|>\omega/v} \frac{\nu_1\nu_2\omega^2}{(\nu_1+\nu_2)^2 q^2 v^2} \tag{6.13}$$

[3]This factorization is particularly useful since χ'' can be written in terms of well known polarization functions. It also provides a simple form amenable to analysis

where $T_0 = vq_0/k_B$. In the degenerate limit $\mu_1, \mu_2 \gg k_B T$ [4] we obtain the energy transfer rate between layers 1 and 2:

$$\mathcal{J}_{12} = \frac{6\zeta(4)}{\hbar^3 v^2} \frac{\nu_1 \nu_2 k_B^4}{(\nu_1+\nu_2)^2} \Big(T_1^4 \ln\frac{T_0}{T_1} - T_2^4 \ln\frac{T_0}{T_2} \Big) \tag{6.14}$$

where $\nu(\mu)$ is the total density of states in each layer, and $k_B T_0 = v(2\pi e^2/\varepsilon)(\nu_1+\nu_2)$. Notably, for equal densities $\mathcal{J}_{12}$ does not depend on the Fermi surface size. For equal densities and small temperature differences between the layers $T_1 \approx T_2$, we obtain the cooling rate

$$\gamma = \frac{1}{C_{\rm el}} \frac{d\mathcal{J}_{12}}{dT} = \frac{9\zeta(4) k_B^2 T^2}{\pi \mu \hbar} \ln\frac{T_0}{T} \tag{6.15}$$

where the heat capacity $C_{\rm el} = \pi^2/3 k_B^2 T \nu(\mu)$ and the density of states $\nu(\mu) = 2\mu/(\pi\hbar^2 v^2)$ for the degenerate limit have been used. Importantly, the rate γ increases as μ goes towards neutrality, but is already quite large for μ away from neutrality. This is completely analogous to intralayer scattering [149, 150]. For typical values $\mu = 100\,\text{meV}$, $T = 300\,\text{K}$, the rate γ is about $10\,\text{ps}^{-1}$, orders of magnitude faster than the electron-lattice cooling rates [11, 12, 26].

6.2.2 Coupled heat transport between layers and E-drag

Vertical energy transfer couples heat transport in the two layers, so that the layer temperatures T_1, T_2 obey

$$\begin{aligned} -\nabla\kappa_1\nabla\delta T_1 + a(\delta T_1 - \delta T_2) + \lambda \delta T_1 &= -\nabla\cdot\mathbf{j}_{q,1} \\ -\nabla\kappa_2\nabla\delta T_2 + a(\delta T_2 - \delta T_1) + \lambda \delta T_2 &= 0 \end{aligned} \tag{6.16}$$

where $a = d\mathcal{J}_{12}/dT$ [see Eq.(6.14)] and λ describes electron-lattice cooling. We consider only a response linear in the applied current, $\mathbf{j}$, neglecting the quadratic joule heating term.

[4] and using the identity

$$\int_{-\infty}^{\infty} \frac{\omega^3}{2} \big(N(\omega) - N(-\omega)\big) d\omega = 12\zeta(4)(k_B T)^4,$$

and recalling that the ultra-violet cutoff arises from screening via $q \ll q_0$

Inverting the coupled linear equations, we find an increase in temperature in layer 2, $\delta T_2(\mathbf{r})$, that is driven by current in layer 1 as

$$\delta T_2(\mathbf{r}) = -\frac{a}{\hat{L_1}\hat{L_2} - a^2}(\mathbf{j}_1 \cdot \nabla)Q[n_1(\mathbf{r}), T], \tag{6.17}$$

where $\mathbf{j}_{q,1}$ is the heat current, Eq.(6.2), where $\hat{L}_i = -\nabla\kappa_i\nabla + a + \lambda$. In what follows we suppress the λ term since electron-lattice cooling is slow. This is a manifestation of the hot carriers in graphene and allows inter-layer vertical energy transfer (described above) to energy transport of the two layers; the above estimates in Eq. 6.15 show that inter-layer energy transfer is far faster than electron-lattice cooling allowing us to ignore λ for the range of chemical potentials of interest. Eq.(6.17) then predicts a value for the interlayer cooling length $\tilde{\ell} = \sqrt{\kappa_1\kappa_2/[(\kappa_1 + \kappa_2)a]}$, which yields a value close to that for the mean free path ℓ. The induced temperature profile, $\delta T_2(\mathbf{r})$, creates thermal gradients that can drive a local thermopower via $\mathbf{E}_2(r) = -(Q[n_2(\mathbf{r})]/T)\nabla\delta T_2$.

Spatial fluctuations in thermopower are governed by density fluctuations via Eq.(6.17). In particular, close to neutrality the local thermopower will exhibit regions of both positive and negative sign, leading to a spatial pattern of the drag resistivity. As discussed above, the correlations between $\delta\mu_1$ and $\delta\mu_2$, Eq.(6.1), lead to a nonzero ensemble-averaged drag resistivity. In the limit $\delta\mu_{1,2} \ll k_BT, \mu_{1,2}$ we write $Q_i(\mathbf{r}) = \langle Q_i(\mathbf{r})\rangle + \frac{\partial Q}{\partial\mu_i}\delta\mu_i(\mathbf{r})$. Passing to Fourier harmonics via $\langle\delta\mu_1(\mathbf{r})\delta\mu_2(\mathbf{r}')\rangle = \sum_{\mathbf{q}} e^{i\mathbf{q}(\mathbf{r}-\mathbf{r}')}\langle\delta\mu_1(-\mathbf{q})\delta\mu_2(\mathbf{q})\rangle$, we obtain Eq.(6.5).

The fact that fluctuating local thermopower, exhibiting both positive and negative signs, does not average to zero is surprising. This happens because the inhomogeneity in heat current and thermopower arise from the same source: electron-hole puddles. E-mechanism resembles mutual drag described by Laikhtman and Solomon [45] in semiconducting heterostructures where doping at contacts produced a similar correlation between Peltier heating/cooling and thermopower. E-mechanism in graphene differs from Ref. [45]

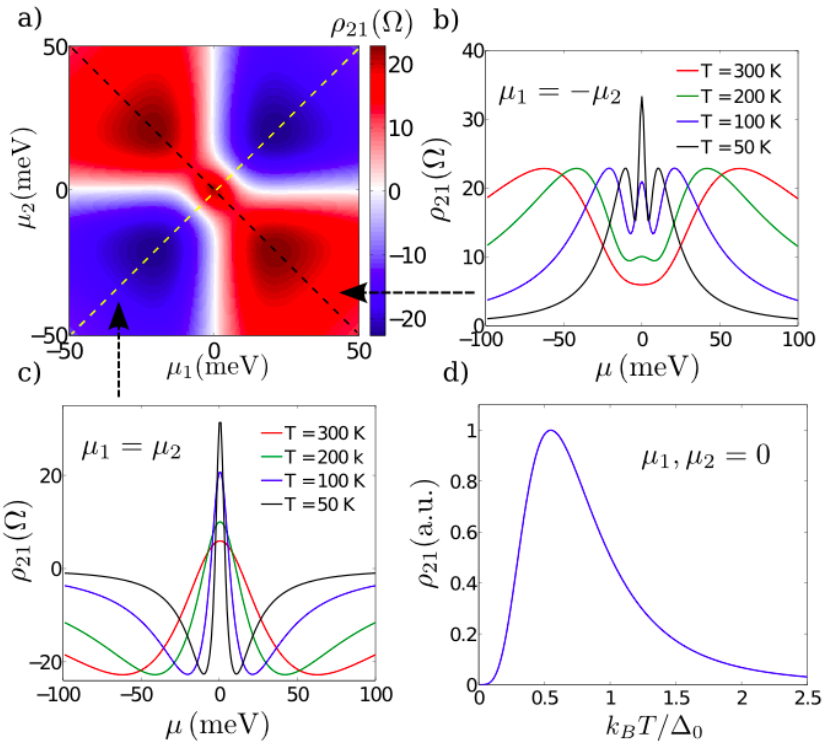

Figure 6.3: (a) Total drag resistivity $\rho_{21}^{(\mathrm{tot})} = \rho_{21}^{(\mathrm{m})} + \rho_{21}^{(\mathrm{e})}$ *vs.* chemical potentials in the two layers, evaluated from Eq.(6.18) and Eq.(6.5) at $T = 100\,\mathrm{K}$, producing a peak at $\mu_{1,2} = 0$ (see text for parameter values used). (b,c) Slices $\mu_1 = \mu_2$ and $\mu_1 = -\mu_2$ at different temperatures. Note a three-peak structure in slice (b) and two sign changes close to CN in (c). (d) Temperature dependence of the peak at $\mu_{1,2} = 0$ in the diffusive regime.

in that density inhomogeneity is intrinsic, occurs throughout the sample (not just at the contacts), and on a far smaller scale.

6.3 Comparing E- and P- Mechanisms

To see how E-mechanism, Eq.(6.5), affects the total experimentally measured drag we need to account for P-mechanism contribution. We use a model that captures the main qualitative features of momentum drag:

$$\rho_{21}^{(\mathrm{m})} = \tilde{\rho}_{21}^{(\mathrm{m})} \frac{h}{e^2} \frac{(k_BT)^2 \mu_1 \mu_2}{(\mu_1^2 + \eta k_B^2 T^2)(\mu_2^2 + \eta k_B^2 T^2)}, \tag{6.18}$$

$\tilde{\rho}_{21}^{(\mathrm{m})} = -1.4\alpha^2/(2\pi\eta^2)$, where $\eta k_B^2 T^2$ describes thermal broadening of the Dirac point. This expression, with the effective interaction strength $\alpha = 0.05$, and $\eta = 6.25$, was obtained by fitting the functional dependence derived in Ref. [140] in the doping region $-10 < \mu/k_{\mathrm{B}}T < 10$.

Combining this with $\rho_{21}^{(\mathrm{e})}$ in Eq.(6.5), we obtain the total drag $\rho_{21}^{(\mathrm{tot})} = \rho_{21}^{(\mathrm{m})} + \rho_{21}^{(\mathrm{e})}$ plotted in Fig.6.3. Here we have used an estimate for thermal conductivity [151]

$$\kappa = b(\mu^2 + \Delta^2(T))/\hbar T, \quad \Delta^2(T) = \Delta_0^2 + \eta(k_B T)^2, \tag{6.19}$$

and assumed Gaussian correlations with average square density fluctuations $\langle \delta\mu^2 \rangle \approx 25\,\mathrm{meV}^2$ and $\xi_{\mathrm{dis}} = 100\,\mathrm{nm}$ [143,144]. Here $\Delta(T)$ accounts for Dirac point broadening by disorder and thermal fluctuations (see Heat Current and Onsager Reciprocity below), and b is a constant of order unity; its exact value does not impact the qualitative features seen in Fig. 6.3. We note that the details of the functional form of the correlator in Eq.(6.1) do not impact the qualitative behavior. The obtained values of total drag are compatible with measured drag resistivities reported in Refs. [44, 141].

The density dependence of total drag plotted in Fig.6.3 (a) can be used to distinguish the two drag mechanisms in experiments. Namely, the peak at zero doping is due to E-mechanism On the slice $\mu_1 = -\mu_2$ (black dashed line) this peak is surrounded by two peaks dominated by the momentum contribution [Fig.6.3(b)]. On the slice $\mu_1 = \mu_2$ (yellow dashed line) the two mechanisms produce contributions of opposite sign, resulting in a double sign change [Fig.6.3(c)]. This provides a clear means of discerning the E-mechanism regime.

The temperature dependence can be estimated as follows. At not too low T such that $\tilde{\ell}, \ell \lesssim \xi_{\mathrm{dis}}$, the sum in Eq.(6.5) yields $\sum_{\mathbf{q}} \langle \delta\mu_1(-\mathbf{q})\delta\mu_2(\mathbf{q})\rangle = \langle \delta\mu_1(\mathbf{r})\delta\mu_2(\mathbf{r}')\rangle_{\mathbf{r}=\mathbf{r}'}$. Using

Eq.(6.3) and κ from Eq.(6.19), we find a non-monotonic T dependence

$$\rho_{21}^{(\mathrm{e})} \propto \frac{T^4}{\left(\Delta_0^2 + \eta(k_B T)^2\right)^3} \langle \delta\mu_1(\mathbf{r}) \delta\mu_2(\mathbf{r}') \rangle_{\mathbf{r}=\mathbf{r}'}, \tag{6.20}$$

This dependence is reminiscent of that reported in Ref. [44] for drag resistance at CN. A similar non-monotonic T dependence arises for in-plane resistivity $\Delta\rho_{\alpha\beta}$. At very low T such that $\ell, \tilde{\ell} \gtrsim \xi_{\mathrm{dis}}$, the sum in Eq.(6.5) is cut at $1/\ell$, giving $\rho_{21} \propto T^8$.

The above analysis can be easily extended to describe the diffusive limit where the elastic mean free path is shorter than the inelastic mean free path, $\ell' < \ell$. Our hydrodynamic approach remains valid in this regime, with the quantity $Q = sT$ where s is the Seebeck coefficient. E-mechanismis still given by Eq.(6.5), with s and κ described by the Mott and Wiedemann-Franz relations:

$$s = \frac{\pi^2}{3e} k_B^2 T \frac{\partial \ln\sigma}{\partial\mu}, \quad e^2\kappa = \frac{\pi^2}{3} k_B^2 T \sigma, \tag{6.21}$$

where σ is the electrical conductivity. Taking σ to vary linearly with carrier density, we find Q that takes on the same qualitative form as Eq.(6.3) in the clean limit. As a result, the qualitative features of $\rho_{21}^{(\mathrm{e})}$ are similar to those found in the clean limit: namely, the avoided crossing of nodal lines, a peak at zero doping, double sign reversal along the diagonal $n_1 = n_2$ and a three-peak structure along the diagonal $n_1 = -n_2$ (Fig.6.3(a,b,c)). The T dependence of $\rho_{21}^{(\mathrm{e})}$ (plotted in Fig. 6.3(d)) is qualitatively similar to the non-monotonic dependence found in the ballistic regime, Eq.(6.20). However, since the Wiedemann-Franz relation gives $\kappa \propto T$ (in contrast to $\kappa \propto 1/T$ in the ballistic regime), at neutrality we find $\rho_{21}^{(\mathrm{e})} \propto T^2$ at lowest T and $\rho_{21}^{(\mathrm{e})} \propto T^{-4}$ at higher $T \gg \Delta$, as shown in Fig.6.3(d). Here, we accounted for Dirac point smearing in the same way as in Eqs.(6.3),(6.19),(6.20).

We note that the effects of energy transport, while being completely generic, are particularly strong in graphene. Since P-mechanism vanishes at CN, whereas E-mechanism

produces a sharp peak in this region, the latter can be easily discerned even at weak inhomogeneity. The peak structure, the sign and the predicted temperature dependence strikingly resemble experiment [44].

6.4 Heat current and Onsager reciprocity

In this section, we provide a detailed derivation of the heat current. While the standard Pelteir coefficient and Mott formula can be used in the diffusive regime, they do not apply in the ballistic regime. Onsager reciprocity allows us to obtain thermopower induced by a temperature imbalance in a Fermi gas in a general form that is applicable in both the ballistic regime, when the mean free path is dominated by electron-electron scattering, and in the diffusive regime, when the mean free path is dominated by elastic scattering by disorder.

We start with recalling Onsager reciprocity for the heat and charge transport. Given charge current $\mathbf{j}$ and heat current $\mathbf{j}_q$ described by [152]

$$-\mathbf{j} = L_{11}\frac{1}{T}\nabla\mu + L_{12}\nabla\frac{1}{T} \tag{6.22}$$

$$\mathbf{j}_q = L_{21}\frac{1}{T}\nabla\mu + L_{22}\nabla\frac{1}{T}, \tag{6.23}$$

the cross-couplings obey the Onsager relation $L_{12} = L_{21}$. Next, we consider the heat current, which can be obtained from the heat transport equation

$$\partial_t(C_{\rm el}T) - \nabla\kappa\nabla T = -\mathbf{j}\cdot\nabla F[n(\mathbf{r}), T], \tag{6.24}$$

where $C_{\rm el}$ is the electron heat capacity, $F[n(\mathbf{r}), T]$ is a function of the carrier density $n(\mathbf{r})$ to be determined later, T is the temperature and κ is the thermal conductivity. From Eq.(6.24) we find that the heat current is

$$\mathbf{j}_{\rm q} = -\kappa\nabla T + \mathbf{j}F[n(\mathbf{r}), T] = -\kappa\nabla T - \sigma F[n(\mathbf{r}), T]\nabla\mu,$$

which gives $L_{21} = -T\sigma F[n(\mathbf{r}), T]$ in Eq.(6.22). Here we have used $\nabla \cdot \mathbf{j} = 0$ and $\mathbf{j} = -\sigma\nabla\mu$, where $\sigma = L_{11}/T$. Using the Onsager relation, the thermopower e.m.f. induced by the temperature gradient equals

$$\mathbf{E} = \frac{F[n(\mathbf{r}), T]}{T}\nabla T. \tag{6.25}$$

This result has general validity irrespective of the transport mechanism specifics, which are manifested through the form of $F[n(\mathbf{r}), T]$.

The functional form of $F[n(\mathbf{r}), T]$ can be obtained by considering the kinetic equation (at steady state)

$$e\mathbf{E} \cdot \nabla_{\mathbf{p}} n(\mathbf{p}, \mathbf{r}) = I_1 + I_2 \tag{6.26}$$

where $e < 0$ is the carrier charge. Here we write the collision integral as a sum of momentum non-conserving and momentum conserving parts, I_1+I_2, corresponding to disorder scattering and electron-electron scattering, respectively. Heat and charge current can be expressed through a steady-state deviation from the equilibrium Fermi distribution, $\delta n = n - n_0$, as follows

$$\mathbf{j}_{\mathrm{q}} = \sum_{\mathbf{p},i} (\epsilon_i - \mu)\mathbf{v}_{\mathbf{p},i}\delta n_i(\mathbf{p}, \mathbf{r}), \quad \mathbf{j} = e\sum_{\mathbf{p},i} \mathbf{v}_{\mathbf{p},i}\delta n_i(\mathbf{p}, \mathbf{r}). \tag{6.27}$$

Here i labels the conduction and valence band states and $\mathbf{v}_{\mathbf{p},i} = \partial\epsilon_i/\partial\mathbf{p}$ and $\epsilon_i = \pm v|\mathbf{p}|$. Below we consider the diffusive and ballistic regimes. In the first case the mean free path is dominated by elastic momentum non-conserving scattering (I_1), in the second regime the mean free path is dominated by inelastic momentum-conserving scattering (I_2).

In the diffusive regime, neglecting I_2 and using the relaxation time approximation for I_1, we find

$$\delta n_{\mathbf{p}} = e\mathbf{E} \cdot \mathbf{v}_{\mathbf{p}}\tau(\epsilon)\partial n_F/\partial\epsilon, \tag{6.28}$$

where τ describes elastic scattering by impurities. This gives the standard expressions for Seebeck and Pelteir coefficient described in Eq. 6.21 so that $F = sT$.

In the ballistic regime, the fastest scattering mechanism comes from the (total) momentum conserving process of electron-electron scattering. As a result, we will neglect all other terms apart from I_2 in Eq. 6.26 and look for distributions, $n(\mathbf{p}, \mathbf{r})$, that give a non-zero particle flow. At a nonzero total current, the non-equilibrium distribution can be written as $n_F(\epsilon - \mathbf{p}\cdot\mathbf{u})$, where the term $\mathbf{p}\cdot\mathbf{u}$ describes the change due to particle flow. This allows us to write δn as

$$\delta n_{\mathbf{p}i} = -\mathbf{u}\cdot\mathbf{p}\frac{\partial n_F}{\partial \epsilon} \tag{6.29}$$

where the Fermi distribution, n_F, has a temperature that may depend on the flow and position. Using Eq. 6.29 and summing over both conduction and valence bands (where $\epsilon_c = v|\mathbf{p}|$ and $\epsilon_v = -v|\mathbf{p}|$, c and v refer to conduction and valence bands respectively) in Eq. 6.27 we obtain

$$\begin{aligned} \mathbf{j}_\mathrm{q} &= \frac{-\mathbf{u}}{2}\int_{-\infty}^{\infty} d\epsilon \nu(\epsilon)\epsilon(\epsilon-\mu)\frac{\partial n_F}{\partial \epsilon}, \\ \mathbf{j} &= \frac{-\mathbf{u}e}{2}\int_{-\infty}^{\infty} d\epsilon \nu(\epsilon)\epsilon\frac{\partial n_F}{\partial \epsilon} \end{aligned} \tag{6.30}$$

where $\nu(\epsilon) = 2|\epsilon|/(\pi\hbar^2 v^2)$ is the total density of states and we take into account that $v_\mathbf{p}\cdot\mathbf{p} = \epsilon$ with ϵ positive (negative) for the conduction (valence) band. Since the function $\epsilon\nu(\epsilon) \propto \epsilon|\epsilon|$, the integral *cannot* be evaluated for arbitrary ratio k_BT/μ. Instead, we analyze the degenerate case, $\mu \gg k_BT$, in which case we use the Sommerfeld expansion to obtain

$$\mathbf{j} = \frac{e}{2}\mu\nu(\mu)\mathbf{u}, \quad \mathbf{j}_\mathrm{q} = 2\zeta(2)\nu(\mu)(k_BT)^2\mathbf{u} \tag{6.31}$$

where we have used the identity $\int_0^\infty \frac{e^x x^2 dx}{(e^x+1)^2} = \zeta(2) = \pi^2/6$. Comparing both expressions to eliminate $\mathbf{u}$ we obtain

$$\mathbf{j}_\mathrm{q} = 4\zeta(2)\frac{(k_BT)^2}{e\mu}\mathbf{j} \tag{6.32}$$

The singularity at $\mu = 0$ is smeared by broadening of the Dirac point due to disorder and

thermal fluctuations. We can account for smearing via

$$\mathbf{j}_{\mathrm{q}} = \frac{4\zeta(2)}{e}\frac{(k_B T)^2 \mu}{\mu^2 + \Delta^2(T)}\mathbf{j}, \quad \Delta^2(T) = \Delta_0^2 + \eta(k_B T)^2 \tag{6.33}$$

where Δ_0 is the Dirac point width parameter. This yields $F[n(\mathbf{r}), T] = Q$, giving Eq. 6.3.

6.5 Summary

In summary, vertical energy transfer in graphene heterostructures has a strong impact on lateral charge transport in the Coulomb drag regime, dominating the drag response at CN. Indeed, the sign and non-monotonic temperature dependence are hallmarks of this new drag mechanism. Importantly, vertical energy transfer being faster than cooling to the lattice, allows interlayer coupled energy transport of electronic systems to play a crucial role in drag. This is another manifestation of the strong influence of hot carriers on the characteristics of graphene.

Drag measurements thus afford a unique probe of energy transfer at the nanoscale, a fundamental process which is not easily amenable to more conventional techniques such as calorimetry, and is key for the physics of strong interactions that occur near neutrality.

Chapter 7

Magnetodrag and Hall Drag in Graphene

Graphene near charge neutrality (CN) hosts an intriguing electron-hole system with unique properties [8, 16, 17, 104, 153–159]. Our understanding of the behavior at CN would greatly benefit from introducing ways to couple the novel neutral modes predicted at CN to charge modes which can be easily probed in transport measurements. There is a long history of employing magnetic field for such a purpose, since transport in charge-neutral plasmas is ultra-sensitive to the presence of external magnetic fields [160].

A new interesting system in which magnetotransport at CN can be probed are atomically thin graphene double layer G/hBN/G structures [44, 134]. As discussed in the previous Chapter, strong Coulomb coupling between adjacent layers in these systems results in strong Coulomb drag, arising when current applied in one (active) layer induces a voltage in the adjacent (passive) layer [44,135,137–141,161–163]. Recent measurements [44] revealed drag resistance that peaks near CN [1] and has dramatic magnetic field dependence, with the peak value increasing by more than an order of magnitude (and changing sign) upon application of a relatively weak B field. Strong magnetic field dependence of drag

[1] An energy-driven drag origin of this peak at $B = 0$ was given in the previous chapter.

has been observed previously in other double layer two-dimensional electron gas (2DEG) heterostructures [164–166], however these experiments were carried out in the quantum Hall regime, whereas the anomalous magnetodrag found in Ref. [44] occurs at classically weak fields $B \lesssim 1\,\mathrm{T}$.

7.1 Energy-driven Mechanism for Magneto and Hall Drag

In this chapter[2], we explain this puzzling $B \neq 0$ behavior in terms of an energy-driven drag mechanism which involves coupled energy and charge transport [162,163] (see Fig.7.1). In the same spirit as that used in the previous chapter, energy transport plays a key role because of fast vertical energy transfer due to interlayer Coulomb coupling in G/hBN/G systems [162] and relatively slow electron-lattice cooling [27,28]. As a result, current applied in one layer can create a spatial temperature gradient for electrons *in both layers*, giving rise to thermoelectric drag voltage. The effect peaks at CN, since thermoelectric response is large close to CN [16,17,159] and diminishes as $1/E_F$ upon doping away from CN [68,75]. We note that unlike E-drag at $B = 0$ (see chapter 6), energy-driven magnetodrag does not require density inhomogeneity and gives a peak at CN; applying B field couples charge and energy modes even at CN without density inhomogeneity. As a result, energy driven magnetodrag is large. Drag arising from this mechanism depends on thermoelectric response and, unlike the conventional momentum drag mechanism, it is insensitive to the electon-electron interaction strength.

Another interesting effect that can be probed in these systems is that of Hall drag. It has long been argued that, at weak coupling, no Hall voltage can arise in the passive

[2]The introduction and 7.1 has been reproduced in part from JCW Song, LS Levitov, Hall Drag and Magnetodrag in Graphene, *Physical Review Letters* **111**, 126601 Copyright (2013) by the American Physical Society. Sec. 7.2 has been reproduced in part with permission from JCW Song, DA Abanin, LS Levitov, Coulomb Drag Mechanisms in Graphene, *NanoLetters* **13** 3631-3637 Copyright 2013 American Chemical Society.

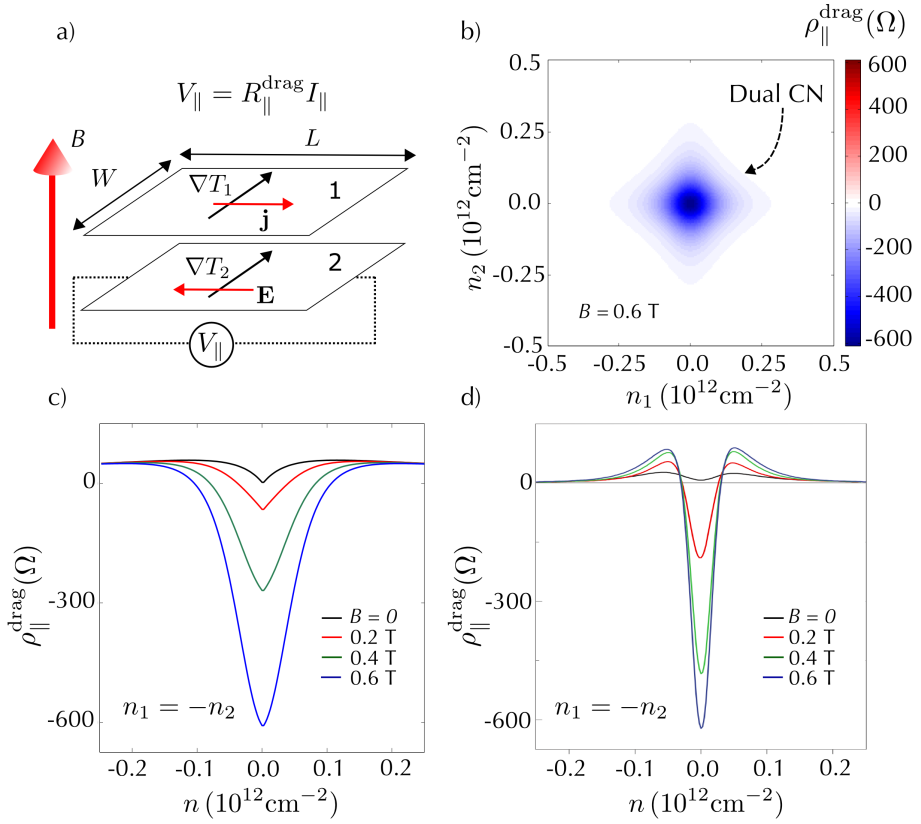

Figure 7.1: Energy-driven magnetodrag in a double layer graphene heterostructure close to CN. (a) Schematic of charge current, temperature gradients, and electric field in the two layers that give rise to a *negative* $\rho_\parallel^{\rm drag}$. (b,c) Magnetodrag resistivity, $\rho_\parallel^{\rm drag}$, obtained from Eqs.(7.11),(7.13). Parameter values: $B = 0.6\,{\rm T}$, $n_0 = 10^{11}\,{\rm cm}^{-2}$, $T = 150\,{\rm K}$, and $\rho_0 = \frac{h}{3e^2}$. The $B = 0$ dependence taken from the model of drag at zero B field [140, 162].(d) Experimentally measured magnetodrag resistivity in G/hBN/G heterostructures, reproduced from Ref. [44] for the same B values as in (c). Application of magnetic field leads to a giant negative drag at CN. Note the similarity between data and theoretically predicted drag density dependence, B dependence, and sign.

layer in the presence of current in the active layer [167, 168]. This is because transferred momentum is parallel to velocity, allowing only a longitudinal "back-current" to develop in the passive layer. As we shall see, a very different behavior arises at strong coupling, owing to the long-range energy currents leading to electron-lattice temperature imbalance. Close to CN, the magnitude of the cross-couplings between charge and energy currents becomes large, producing a finite Hall drag, $V_H = R_H^{\rm drag} I_\parallel$.

As we will see, energy currents result in Hall and magnetodrag resistances, $R_H^{\rm drag}$ and $R_\parallel^{\rm drag}$, that are large and peak near CN, see Fig.7.1 and Fig.7.2. These large values arise even for classically weak fields $B \sim 0.1\,\mathrm{T}$, exceeding by two orders of magnitude the values found previously in GaAs/AlGaAs 2DEG heterostructures [164–166] at similar fields. The mechanism based on coupled energy and charge transport predicts large and negative drag at CN that matches recent experiments (see Fig.7.1c,d). Our mechanism naturally leads to Hall drag because vertical energy transfer between layers does not discriminate between longitudinal and transverse heat currents since temperature profile is a scalar field. This stands in contrast to conventional momentum driven drag, where momentum transfer is parallel to the applied current [167, 168].

7.1.1 Coupled Charge and Heat Transport in a Magnetic Field

Heat current and an electric field, induced by charge current and temperature gradients, are coupled via the thermoelectric effect altered by the B field,

$$\mathbf{j_q} = \underline{\mathbf{Q}}\mathbf{j}, \quad \mathbf{E} = \underline{\mathbf{Q}}\frac{\nabla T}{T}. \tag{7.1}$$

Here $\underline{\mathbf{Q}}$ is a 2×2 matrix, of which diagonal components describe the Peltier and Thompson effects, and off-diagonal components describe the Nernst-Ettingshausen effect. Onsager reciprocity requires that $\underline{\mathbf{Q}}$ in both the expressions for $\mathbf{j_q}$ and $\mathbf{E}$ are the same [see analysis

following Eq.(7.9)]. As an example of how our mechanism produces drag consider the Hall bar geometry, see Fig.7.1. When a longitudinal charge current is applied in the active layer (for $B \neq 0$) a transverse (Ettingshausen) heat current develops in both layers through efficient vertical energy transfer. Nernst voltage in the passive layer results in a longitudinal magnetodrag of a *negative sign*.

To obtain the electric field in layer 2 induced by current applied in layer 1, we first need to understand the coupling of temperature profiles $T_{1,2}(\mathbf{r})$ in the two layers. Energy transport in the system can be described by

$$\begin{aligned} -\nabla \underline{\kappa}_1 \nabla \delta T_1 + a(\delta T_1 - \delta T_2) + \lambda \delta T_1 &= -\nabla \cdot \left(\underline{\mathbf{Q}}^{(1)} \mathbf{j}\right) \\ -\nabla \underline{\kappa}_2 \nabla \delta T_2 + a(\delta T_2 - \delta T_1) + \lambda \delta T_2 &= 0 \end{aligned} \tag{7.2}$$

with a the energy transfer rate between the two layers [162], λ the electron-lattice cooling rate, and $\delta T_i = T_i - T_0$ (here T_0 is the lattice temperature, equal for both layers).

Here we focus on a Hall-bar geometry of two parallel rectangular layers of dimensions $L \times W$, $L \gg W$, pictured in Fig.7.1a. For the sake of simplicity, we treat the electric and heat currents as independent of the x coordinate along the bar. In layer 1, current is injected at $x = -L/2$ and drained at $x = L/2$. In layer 2, the Hall drag voltage arising across the device, V_H and the longitudinal drag voltage, $V_\parallel$, are evaluated as

$$V_H = \int_{-W/2}^{W/2} E_y^{(2)} dy, \quad V_\parallel = \frac{L}{W} \int_{-W/2}^{W/2} E_x^{(2)} dy \tag{7.3}$$

The electric and thermal variables may depend on the transverse coordinate y, see below.

Boundary conditions for a Hall bar require electric current being tangential to the side boundaries, $y = \pm W/2$, and zero temperature imbalance at the ends, $x = \pm L/2$, reflecting that the current and voltage contacts act as ideal heat sinks. The electric current parallel to the boundaries $y = \pm W/2$ gives rise to the Ettingshausen heat current that may have a component transverse to the Hall bar. The divergence of this heat current, appearing

on the right hand side of Eq.(7.2), acts as an effective boundary delta function source in the heat transport equations. Boundary conditions can profoundly influence the symmetry of the resultant drag resistivity, see below.

We consider the case of a spatially uniform $\underline{\mathbf{Q}}$ in both layers. The ideal heat sinks at $x = \pm L/2$ mean that no temperature imbalance develops in the x-direction (except for some "fringing" heat currents near the contacts which give a contribution small in $W/L \ll 1$, which we will ignore in the following discussion). Since no temperature gradients are sustained in the x-direction far from the ends, we can reduce our problem Eq.(7.2) to a quasi-1D problem with temperature profiles that only depends on the y-coordinate. As a result, the only heat source arises from the Ettingshausen effect $\underline{\mathbf{Q}}^{(1)}\mathbf{j} = (Q^{(1)}_{yx} j)\hat{\mathbf{y}}$.

To describe transport in the presence of such a source, we will expand temperature variables in both layers in a suitable orthonormal set of functions. Here it will be convenient to use eigenstates of the operator ∂_y^2 with zero Neumann boundary conditions at $y = \pm W/2$, given by

$$u_n(y) = A\cos\left(\frac{2\pi n}{W}y\right), \quad v_n(y) = A\sin\left(\frac{2\pi(n+\frac{1}{2})}{W}y\right),$$

$A = (2/W)^{1/2}$, $n = 0, 1, 2...$ From the symmetry of the source in Eq.(7.2) we expect $\delta T_{1,2}(y)$ to be odd in y. Thus only the functions $v_n(y)$ are relevant, giving

$$\delta T_{1,2}(y) = \sum_{q_n} \delta\tilde{T}_{1,2}(q_n) A \sin q_n y, \quad q_n = \frac{2\pi(n+\frac{1}{2})}{W}.$$

For each n we obtain a pair of algebraic equations

$$\begin{aligned} q_n^2\kappa_1\delta\tilde{T}_1 + a(\delta\tilde{T}_1 - \delta\tilde{T}_2) + \lambda\delta\tilde{T}_1 &= F_n \\ q_n^2\kappa_2\delta\tilde{T}_2 + a(\delta\tilde{T}_2 - \delta\tilde{T}_1) + \lambda\delta\tilde{T}_2 &= 0 \end{aligned} \tag{7.4}$$

where $\kappa_{1,2} = \kappa_{xx}^{(1,2)}$ and $F_n = 2A(-1)^n Q^{(1)}_{yx} j$. Solving Eq.(7.4), we find the temperature

profile in layer 2:

$$\delta T_2(y) = \sum_{n\geq 0} \frac{aF_n}{L_1(q_n)L_2(q_n) - a^2} v_n(y), \tag{7.5}$$

where $L_i(q_n) = \kappa_i q_n^2 + a + \lambda$ $(i = 1, 2)$. Since electron-lattice cooling is very slow [27, 28], with the corresponding electron-lattice cooling length values in excess of few microns, we will suppress λ in what follows. Because the boundaries in the transverse (y-direction) are free, a finite temperature imbalance between the edges can arise, given by $\Delta T = \delta T_2(y = W/2) - \delta T_2(y = -W/2)$. We find

$$\Delta T = 4A^2 \sum_{n\geq 0} \frac{aQ_{yx}^{(1)} j}{L_1 L_2 - a^2} = \frac{8}{W\tilde{\kappa}} \sum_{n\geq 0} \frac{Q_{yx}^{(1)} j}{q_n^2(1 + \xi_{\rm inter}^2 q_n^2)}, \tag{7.6}$$

where we defined $\tilde{\kappa} = \kappa_1 + \kappa_2$ and a length scale $\xi_{\rm inter} = \sqrt{\kappa_1\kappa_2/a\tilde{\kappa}}$ (interlayer thermal equilibration length). We evaluate the sum using the identity $\sum_{n=0}^{\infty} \frac{1}{(n+\frac{1}{2})^4 + c^2(n+\frac{1}{2})^2} = \frac{\pi^2}{2c^2}\left(1 - \frac{\tanh \pi c}{\pi c}\right)$ to obtain

$$\Delta T = \frac{WQ_{yx}^{(1)} j}{\tilde{\kappa}} G(\xi), \quad G(\xi_{\rm inter}) = 1 - \frac{2\xi_{\rm inter}}{W} \tanh\left(\frac{W}{2\xi_{\rm inter}}\right). \tag{7.7}$$

Connecting ΔT with the drag voltage, and in particular determining its sign, requires taking full account of Onsager reciprocity. This analysis is presented below.

In the same way that the applied charge current, $\mathbf{j}$, in layer 1 causes a heat current (Peltier/Ettingshausen), a temperature imbalance in layer 2, ΔT, can sustain voltage drops across the sample (Thermopower/Nernst). These two effects are related by Onsager reciprocity constraints. The cross couplings in the coupled energy and charge transport equations [152] arise from

$$\left(\begin{array}{c} -\mathbf{j} \\ \mathbf{j_q} \end{array}\right) = \left(\begin{array}{c|c} e\underline{\mathbf{L}}_{11}/T & e\underline{\mathbf{L}}_{12} \\ \hline \underline{\mathbf{L}}_{21}/T & \underline{\mathbf{L}}_{22} \end{array}\right) \left(\begin{array}{c} \nabla\mu \\ \nabla\frac{1}{T} \end{array}\right) \tag{7.8}$$

where $\underline{\mathbf{L}}$ are 2×2 matrices and e is the carrier charge. In this notation, the electrical conductivity equals $\underline{\sigma} = e^2\underline{\mathbf{L}}_{11}/T$, and thermal conductivity is $\underline{\kappa} = \underline{\mathbf{L}}_{22}/T^2$. Comparing to the heat current due to an applied charge current, Eq.(7.1), we identify $\underline{\mathbf{L}}_{21} = -e\underline{\mathbf{Q}}\underline{\mathbf{L}}_{11}$.

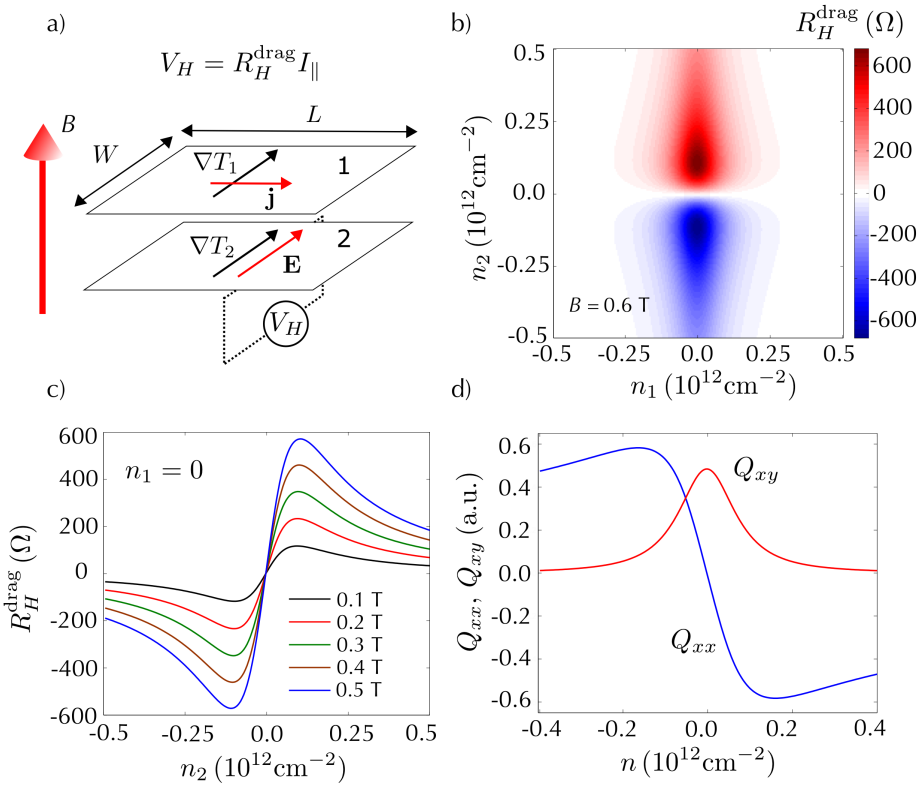

Figure 7.2: Energy-driven Hall drag in double layer graphene heterostructures (a) Schematic of charge current, temperature gradients, and electric field in the two layers of a Hall bar that produces Hall drag. (b,c) Density dependence of Hall drag resistance, predicted from Eqs.(7.11),(7.13) for the same parameter values as in Fig.7.1. (d) Density dependence of Q_{xx}, Q_{xy}, see text.

Onsager reciprocity demands that the cross-couplings obey $\underline{\mathbf{L}}_{12}(B) = \underline{\mathbf{L}}_{21}^{\mathrm{T}}(-B)$ where B is the applied magnetic field (note the transposed matrix). In an isotropic system the off-diagonal components of $\underline{\mathbf{L}}$ obey $\underline{\mathbf{L}}^{(xy)}(B) = \underline{\mathbf{L}}^{(yx)}(-B)$. As a result, Onsager reciprocity reduces to

$$\underline{\mathbf{L}}_{12}(B) = \underline{\mathbf{L}}_{21}(B) \tag{7.9}$$

in an isotropic system. Applying Eq. 7.9 to Eq. 7.8 in an open circuit, we find $\mathbf{E} = -e^{-1}\nabla\mu = T^{-1}\underline{\mathbf{L}}_{11}^{-1}\underline{\mathbf{Q}}\underline{\mathbf{L}}_{11}\nabla T$. For an isotropic system $\underline{\mathbf{Q}} = Q_{xx}\mathbf{1} + iQ_{xy}\sigma_2$, $\underline{\mathbf{L}} = L_{xx}\mathbf{1} + iL_{xy}\sigma_2$, so that $[\underline{\mathbf{Q}}, \underline{\mathbf{L}}] = 0$, which gives Eq.(7.1). $\sigma_{1,2}$ are the x, y Pauli matrices.

Several different regimes arise depending on the relation between the interlayer thermal equilibration length ξ_{inter} and the bar width W. Using Eq.(7.3) and Eq.(7.1) we obtain

$$\begin{pmatrix} V_{\parallel} \\ V_H \end{pmatrix} = \begin{pmatrix} R_{\parallel}^{\mathrm{drag}} & -R_H^{\mathrm{drag}} \\ R_H^{\mathrm{drag}} & R_{\parallel}^{\mathrm{drag}} \end{pmatrix} \begin{pmatrix} I_{\parallel} \\ 0 \end{pmatrix}, \tag{7.10}$$

giving the magnetodrag and Hall drag resistance values

$$R_H^{\mathrm{drag}} = \frac{-G(\xi_{\mathrm{inter}})}{T\tilde{\kappa}} Q_{xy}^{(1)} Q_{xx}^{(2)}, \quad R_{\parallel}^{\mathrm{drag}} = \frac{-LG(\xi_{\mathrm{inter}})}{WT\tilde{\kappa}} Q_{xy}^{(1)} Q_{xy}^{(2)}, \tag{7.11}$$

where we used $Q_{xx} = Q_{yy}$ and $Q_{xy} = -Q_{yx}$ for an isotropic system. For a narrow bar (or, slow interlayer equilibration), we have $\xi_{\mathrm{inter}}/W \gg 1$ and $G \to 0$, giving vanishingly small $R_{H,\parallel}^{\mathrm{drag}}$. For a wide bar (or, fast interlayer equilibration) we have $G \to 1$ so that $R_{H,\parallel}^{\mathrm{drag}}$ saturates to a universal value independent of the interlayer cooling rate. For typical device parameters, we estimate $\xi_{\mathrm{inter}} \approx 40\,\mathrm{nm}$ at $T = 300\,\mathrm{K}$ [162]. Since L, W are a few mircons for typical graphene devices, we expect them to be firmly in the $G = 1$ regime, with the Hall drag and magnetodrag attaining universal values independent of the electron-electron interaction strength.

7.1.2 Density and B Field dependence

To describe the density and B field dependence, we use a simple model for $\underline{\mathbf{Q}}$. Measurements indicate [16, 17] that thermopower and the Nernst effect in graphene are well described by the Mott formula [169], giving

$$\underline{\mathbf{Q}} = \frac{\pi^2}{3e} k_B^2 T^2 \underline{\rho} \frac{\partial[\underline{\rho}^{-1}]}{\partial \mu}, \quad \underline{\rho} = \begin{pmatrix} \rho_{\parallel} & \rho_H \\ -\rho_H & \rho_{\parallel} \end{pmatrix}, \tag{7.12}$$

with ρ the resistivity, $e < 0$ the electron charge, and μ the chemical potential. We use a simple phenomenological model [170] relevant for classically weak B fields:

$$\rho_{\parallel} = \frac{\rho_0}{\sqrt{1 + n^2/n_0^2}}, \quad \rho_H = \frac{-Bn}{e(n^2 + n_0^2)}, \tag{7.13}$$

where ρ_0 is the resistivity peak value at the Dirac point, n is the carrier density, and parameter n_0 accounts for broadening of the Dirac point due to disorder. We account for disorder broadening of the density of states, $dn/d\mu = (n^2 + n_0^2)^{1/4} \left(2/(\pi \hbar^2 v_F^2)\right)^{1/2}$.

From Eqs.(7.11),(7.12),(7.13) and the Wiedemann-Franz relation for κ, we obtain $\rho_{\parallel}^{\mathrm{drag}} = (W/L) R_{\parallel}^{\mathrm{drag}}$ and R_H^{drag} (see Fig.7.1b,c and Fig.7.2b,c, respectively). In that, we used the parameter values $n_0 = 10^{11}\,\mathrm{cm}^{-2}$, $\rho_0 = \frac{h}{3e^2}$, and a representative temperature, $T = 150\,\mathrm{K}$. These values match device characteristics (disorder broadening, n_0, and peak resistivity, ρ_0) described in Ref. [44]. As a sanity check, we plot the components of $\underline{\mathbf{Q}}$ (in Fig.7.2d) which show the behavior near CN matching thermopower and Nernst effects measured in graphene [16, 17].

Analyzing magnetodrag, we find that $\rho_{\parallel}^{\mathrm{drag}}$ peaks at dual CN, taking on large and *negative* values (Fig.7.1b,c). The Magnetodrag peak exhibits a steep B dependence, $\rho_{\parallel,\mathrm{peak}}^{\mathrm{drag}} \propto -B^2$, bearing a striking resemblance to measurements reproduced in Fig.7.1d. In particular, our model explains the negative sign of the measured magnetodrag.

Hall drag is large and sign-changing (see Fig.7.2b,c), taking on values consistent

with measurements [171]. Interestingly, the map in Fig.7.2b indicates that the sign of $R_H^{\rm drag}$ is controlled solely by carrier density in layer 2, breaking the $n_1 \leftrightarrow n_2$ symmetry between layers. This behavior does not contradict Onsager reciprocity, it arises as a consequence of the asymmetric boundary conditions for the Hall bar: free boundary at $y = \pm W/2$ and ideal heat sinks at the ends, $\delta T(x = \pm L/2) = 0$. This allows for finite temperature gradients to be sustained across the bar but not along the bar, see Fig.7.2a.

For other geometries, the temperature gradient can be obtained by balancing the heat flux due to thermal conductivity against the net heat flux in the two layers, $(\kappa_1 + \kappa_2)\nabla\delta T = \underline{\mathbf{D}}\,\underline{\mathbf{Q}}^{(1)}\mathbf{j}_1$. The quantity $\underline{\mathbf{D}}$ can in principle be obtained by solving heat transport equations. Adopting the same approach as above, we find a magneto and Hall-drag resistivity

$$\underline{\rho}^{\rm drag} = \frac{1}{T\tilde{\kappa}}\underline{\mathbf{Q}}^{(2)}\underline{\mathbf{D}}\,\underline{\mathbf{Q}}^{(1)}, \quad \mathbf{E}_2 = \underline{\rho}^{\rm drag}\mathbf{j}_1. \tag{7.14}$$

For isotropic heat flow, $\underline{\mathbf{D}} = \mathbf{1}$. In this case, since $\underline{\mathbf{Q}}^{(1)}$ and $\underline{\mathbf{Q}}^{(2)}$ commute, the resulting drag is layer-symmetric, $n_1 \leftrightarrow n_2$ (see 7.2). In particular, Hall drag for $\underline{\mathbf{D}} = \mathbf{1}$ vanishes on the diagonal $n_1 = -n_2$. In contrast, for anisotropic heat flow, such as that discussed above, we expect a generic tensor $\underline{\mathbf{D}} \neq \mathbf{1}$ and thus no layer symmetry.

We wish to clarify, in connection to recent measurements, [171] that layer symmetry $n_1 \leftrightarrow n_2$ implies a swap of current and voltage contacts. Layer symmetry, which implies $\underline{\mathbf{D}} = \mathbf{1}$ in Eq.(7.14), will therefore only hold for Hall bars equipped with wide voltage contacts, for which the contact and the bar widths are comparable. This is indeed the case for the cross-shaped devices used in Ref. [44]. However it is not the case for a Hall bar with *noninvasive* voltage probes which are much narrower than the bar width, as assumed in our analysis above. Noninvasive probes, which have little effect on temperature distribution in the electron system, translate into $\underline{\mathbf{D}} \neq \mathbf{1}$ and no layer symmetry.

In summary, magnetic field has dramatic effect on drag at CN because it induces

strong coupling between neutral and charge modes, which are completely decoupled in the absence of magnetic field in a uniform system. Field-induced mode coupling leads to giant drag that dwarfs the conventional momentum drag contribution as well as a remnant drag due to spatial inhomogeneity (discussed in chapter 6 and Ref. [162]). Our estimates indicate that these two contributions are orders of magnitude smaller than the predicted magnetodrag, which also has an opposite sign. The giant magnetodrag and Hall drag values attained at classically weak magnetic fields, along with the unique density dependence and sign, make these effects easy to identify in experiment. The predicted magnetodrag is in good agreement with findings in Ref. [44]. Magnetic field, coupled with drag measurements at CN, provides a unique tool for probing the neutral modes in graphene.

7.2 Comparing E- and P- Mechanisms for Magneto/Hall Drag

Coulomb drag has a rich history and occurs despite the lack of particle exchange between layers and provides one of the most sensitive probes of interactions in low-dimensional systems. Indeed, as we have seen, long-range Coulomb interactions have long been known to result in a kind of spooky action between adjacent electrically isolated electron systems arising when current applied in one (active) layer induces voltage in the second (passive) layer, Fig. 7.3a. Coulomb drag was extensively studied in GaAs quantum wells, [46, 142] where the observations were successfully interpreted in terms of the momentum drag mechanism [167,172–174] (hereafter referred to as "P-mechanism"), in which interlayer electron-electron scattering mediated by long-range Coulomb interaction transfers momentum from the active layer to the passive layer.

In the previous section, we focussed on the E-mechanism and established how it gives rise to magneto and Hall drag in graphene. In this section, we compare it with the P-

mechanism. As we will see, E and P mechanisms give qualitatively different drag responses.

While on a microscopic level the P and E mechanisms both arise from the same electron-electron interactions, the two contributions to drag are associated with very different physical effects: interlayer momentum transfer vs. interlayer energy transfer and long-range lateral energy transport coupled to charge flow. Accordingly, these effects develop on very different lengthscales. For the P-mechanism the characteristic lengthscales are on the order of the Fermi wavelength, which makes this mechanism essentially local. In contrast, the E-mechanism originates from lateral energy transport in the electronic system which is highly nonlocal. As a result, the two mechanisms have been described by very different approaches, microscopic for P-mechanism [135–140] and hydrodynamical for E-mechanism (previous chapter and previous section in this chapter also Ref. [162,175]) which reflects the difference in the characteristic lengthscales.

Here we adopt a different strategy and develop a unified framework capable of describing these mechanisms on an equal footing. To tackle the different lengthscales relevant for the E and P mechanisms, a suitable multiscale framework is needed. This framework should also account for the peculiar features of particle and hole dynamics near the Dirac point: Electric currents carried by electrons and holes in the same direction generate momentum flow in opposite directions. [157,158] The same is true for energy flow due to particle and hole currents. In the presence of a magnetic field, the opposite sign of the Lorentz force on electrons and holes makes charge currents strongly coupled to neutral (energy) currents, resulting in a giant Nernst/Ettingshausen effect. [16, 17, 150, 159]

As we will show, this rich behavior is conveniently captured by a simple two-fluid model. In this model, carriers in the conduction and valence bands are described as separate subsystems coupled by mutual drag, originating from carrier-carrier scattering. As we will see, P-mechanism and E-mechanism drag, for both $B = 0$ and $B \neq 0$, can

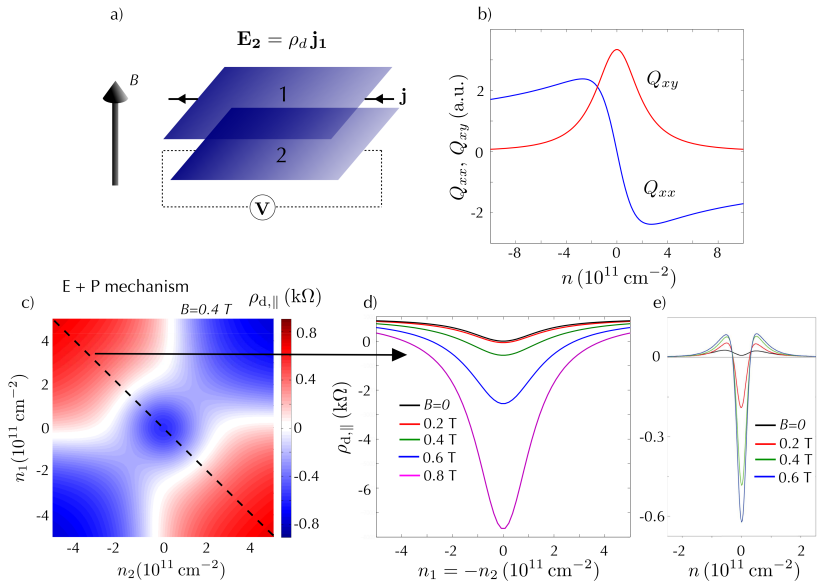

Figure 7.3: Two adjacent layers of graphene can exhibit a drag resistivity, ρ_{d}, that features giant enhancement near the double neutrality point (DNP). (a) Schematic of double-layer graphene heterostructure and drag measurement. (b) Thermoelectric coupling $Q_{xx} = \mathrm{Re(Q)}$, $Q_{xy} = -\mathrm{Im}(Q)$, Eq.(7.31), peaks close to the Dirac point. (b) Magnetodrag, $\rho_{\mathrm{d},\parallel}$, obtained by summing E-mechanism and P-mechanism contributions with weighting factor $\beta = 0.3$, see Eqs.(7.27),(7.34),(7.33) [here $B = 0.4\,\mathrm{T}$]. P-mechanism dominates far from DNP, whereas E-mechanism dominates close to DNP creating a unique configuration of nodal lines. The large negative peak of magnetodrag at DNP is a hallmark of the energy-transport mechanism. Parameters used are same as in Fig. 7.4. (d) Line trace ($n_1 = -n_2$) along dashed line in panel (c) for various values of B. (e) Experimental measurement of $\rho_{\mathrm{d},\parallel}$ from Ref. [44] displaying the behavior at DNP similar to (d).

be obtained from the same two-fluid model, allowing us to analyze these contributions to drag even-handedly. A model of this type was developed a while ago by Gantmakher and Levinson [176] to describe magnetotransport in charge-compensated conductors, and in particular the anomalies in magnetoresistance and Hall resistance arising at nearly equal electron and hole densities. A similar model was used to describe magnetotransport near CN in graphene in Ref. [177]. As we will see, this model can successfully account for the strong influence of magnetic field on drag near DNP observed in Ref. [44].

The two-fluid model predicts density and magnetic field dependence which is distinct for the P and E contributions (see Fig. 7.4). The E contribution features a large peak at DNP, whereas the P contribution is small near DNP and large away from DNP. The peak in the E contribution is sharply enhanced by magnetic field, whereas the P contribution does not show strong field dependence (see Fig. 7.4). Overall, the density plots for P and E contributions look similar up to an overall sign reversal. This behavior makes it easy to distinguish these contributions in experiment.

We find that the magnitude of the E contribution can exceed the P contribution as is evident near DNP. Indeed, adding the two contributions up in Fig. 7.3c,d produces a distinct density dependence of $\rho_{\mathrm{d},\parallel}$ (Fig. 7.3c). While the exact arrangement of nodal lines can somewhat depend on the parameters chosen, the qualitative features - E-mechanism dominates near DNP (negative $\rho_{\mathrm{d},\parallel}$), and P-mechanism dominates far away from DNP - are robust.

The general reason for the relative smallness of the P contribution can be understood as follows. The two-fluid model describes coupling between carriers of different types via the mutual drag coefficient η, see Eq.(7.19). The dependence of η on the interaction strength $\alpha_0 = e^2/\hbar v$ can be obtained [177] by matching the dependence of conductivity *vs.*

α_0 found in Refs. [157, 158]. This gives a general relation of the form

$$\eta = F(\alpha)\hbar, \quad \alpha = \frac{\alpha_0}{\kappa_0 + \frac{\pi}{8} N \alpha_0}, \tag{7.15}$$

where α is the RPA-screened interaction, N is spin/valley degeneracy and κ_0 is the dielectric constant for the substrate. Since $N = 8$ for the double layer system, the factor in the denominator is ~ 10 ($\kappa_0 \approx 4$ for a BN substrate). A ten-fold reduction of the bare value $\alpha_0 \approx 2.4$ yields a small value of RPA-screened interaction, $\alpha \approx 0.25$. The function F admits a power series expansion in α, arising from the solution of the quantum Boltzmann equation, [157, 158] with the leading term being $\alpha^2 \sim 0.06$ (which corresponds to the two-particle Born scattering cross-section). This leads to a weak mutual drag, $\eta \ll \hbar$.

Crucially, the E contribution to drag remains unaffected by the small values of η so long as the interlayer thermalization occurs faster than the electron-lattice relaxation. This is the case in graphene, since electron-lattice cooling in this material is dominated by acoustic phonons, giving a slow electron-lattice cooling rate in a wide range of temperatures. [26–28] As a result, as discussed in more detail in Sec.7.2.3, the drag originating from E-mechanism takes on a "universal value" which shows little dependence on the interlayer scattering rate. The relative strength of the P and E contributions to magnetodrag at DNP, estimated below, can be characterized by

$$\rho_{\mathrm{d},\parallel}^{(\mathrm{P})} \approx -\left(\frac{0.6}{\beta}\frac{\sigma\eta}{e^2}\right)\rho_{\mathrm{d},\parallel}^{(\mathrm{E})} \tag{7.16}$$

where σ is the conductivity at charge neutrality and $\beta \sim 1$ is a factor describing temperature gradient buildup in response to energy flow in the system. Since $\sigma \sim 4e^2/h$ whereas η is much smaller than $\hbar$, the factor in parenthesis is much smaller than unity. The smallness of the P contribution, while being quite general, is not entirely universal. In particular, it does not hold far from DNP, where E-mechanism is small (see Fig.7.4). It also does not hold at elevated temperatures when electron-lattice cooling length becomes small compared

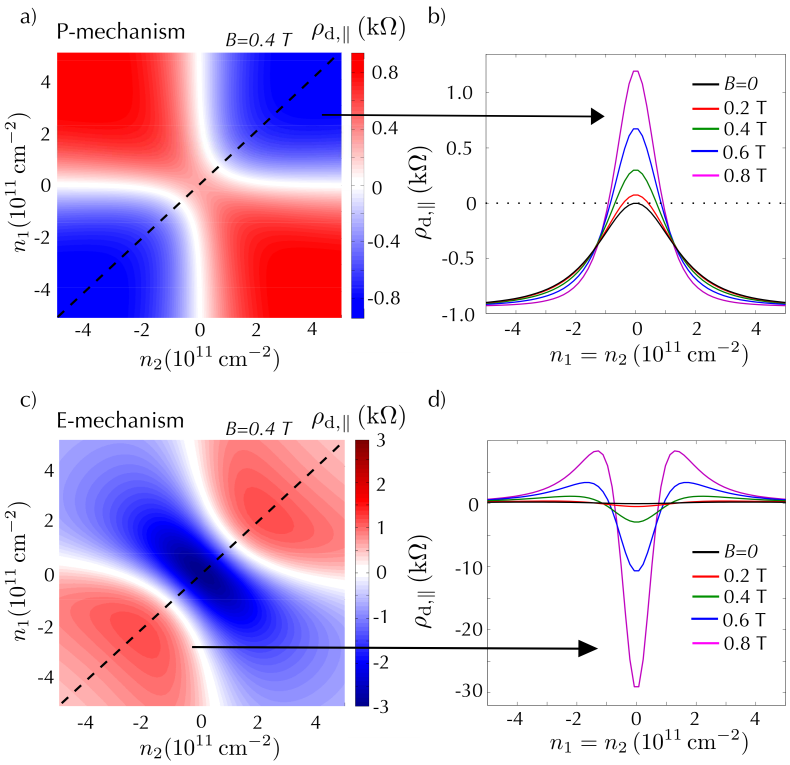

Figure 7.4: Coulomb drag originating from P and E mechanisms for different values of applied magnetic field. (a) Magnetodrag density dependence for P-mechanism, obtained from Eq.(7.27) for $B = 0.4\,\mathrm{T}$, $\eta = 0.23\hbar$. Note that drag resistance $\rho_{\mathrm{d},\parallel}$ is weaker at DNP than away from DNP, and has positive sign at DNP. (b) Line traces along dashed line in (a) for various values of B. (c) Magnetodrag density dependence for E-mechanism, obtained from Eq.(7.34) at $B = 0.4\,\mathrm{T}$, $\eta = 0$ and $\beta = 1$. Note a large negative peak at DNP which is strongly enhanced by B field, a behavior distinct from that for P-mechanism. (d) Line traces along dashed line in (c) for various values of B. Disorder broadening of DP of width $\Delta \approx 200\,\mathrm{K}$ and $T = 200\,\mathrm{K}$ were used here as well as in Figs. 7.3, 7.5 (see text).

to system size, $\xi \lesssim W$, leading to small values $\beta \ll 1$ which suppress the E contribution, see Eq.(7.39). However, near DNP and at not too fast electron-lattice cooling, we expect E-mechanism to overwhelm P-mechanism.

7.2.1 The Two-fluid model

To describe transport near the Dirac point, it is crucial to account for the contributions of both electrons and holes. This can be done by employing the quantum kinetic equation approach of Refs. [157, 158, 176, 177]. For a spatially uniform system, we have

$$q_{e(h)} \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \frac{\partial f_{e(h)}(\mathbf{p})}{\partial \mathbf{p}} = I[f_e(\mathbf{p}), f_h(\mathbf{p})], \tag{7.17}$$

where $f_{e(h)}(\mathbf{p})$ is the distribution function for electrons and holes, and $q_e = -q_h = e$. The collision integral I describes momentum relaxation due to two-particle collisions and scattering by disorder. The approach based on Eq.(7.17) is valid in the quasiclassical regime, when particle mean free paths are long compared to wavelength. This is true for weak disorder and carrier-carrier scattering.

The kinetic equation (7.17) can be solved analytically in the limit of small α. [157, 158] Rather than pursuing this route, we will adopt a two-fluid approximation used in Refs. [176,177] which is particularly well suited for analyzing magnetotransport. In the two-fluid approach, transport coefficients can be obtained from the balance of the net momentum for different groups of carriers, electrons and holes, taken to be moving independently. We use a simple ansatz for particle distribution function,

$$f_{e(h)}(\mathbf{p}) = \frac{1}{e^{(\epsilon_{\mathbf{p}} - \mathbf{p}\mathbf{a}_{e(h)} - \mu_{e(h)})/k_{\mathrm{B}}T} + 1}, \quad \epsilon_{\mathbf{p}} = v_0|\mathbf{p}|, \tag{7.18}$$

where $\mu_e = -\mu_h$ are the chemical potentials of electrons and holes. The quantities $\mathbf{a}_e$ and $\mathbf{a}_h$, which have the dimension of velocity, are introduced to describe a current-carrying state. This ansatz corresponds to a uniform motion of the electron and hole subsystems, such that

the collision integral for the e-e and h-h processes vanishes (as follows from the explicit form of the collision integral given in Ref. [158]). Thus only the e-h collisions contribute to momentum relaxation, resulting in mutual drag between the e and h subsystems.

In the following analysis we do not account for possible temperature imbalance between electron and hole subsystems since fast e-e and e-h collisions quickly establish thermal equilibrium locally in space. As we will see below, spatial temperature variation across the system becomes essential in the regime dominated by energy transport. We will treat this regime in Sec.7.2.3.

Eq.(7.17) yields coupled equations for ensemble-averaged velocities and momenta of different groups of carriers, described by the distribution (7.18):

$$-q_i\left(\mathbf{E}_i+\mathbf{V}_i\times\mathbf{B}\right)=-\frac{\mathbf{P}_i}{\tau_i}-\eta\sum_{i'}n_{i'}(\mathbf{V}_i-\mathbf{V}_{i'}), \tag{7.19}$$

where $i,i'=1,2,3,4$ label the e and h subsystems in the two layers. The ensemble-averaged scattering times τ_i, the carrier densities n_i, and the electron-hole drag coefficient η, describing collisions between electrons and holes, are specified below. The electric field $\mathbf{E}_i$ is the same for electrons and holes in one layer, but is in general different in different layers.

The quantities $\mathbf{V}_i$, $\mathbf{P}_i$ are proportional to each other, $\mathbf{P}_i=m_i\mathbf{V}_i$. Here the "effective mass" is obtained by averaging over the distribution of carriers, as described in 7.2.4. The integrals over $\mathbf{p}$, evaluated numerically, give the effective mass as a function of T and μ. At charge neutrality, setting $\mu_{e(h)}=0$, we find

$$m=\frac{9\zeta(3)}{2\zeta(2)}\frac{k_{\mathrm{B}}T}{v_0^2}\approx 3.288\frac{k_{\mathrm{B}}T}{v_0^2}. \tag{7.20}$$

At high doping, $\mu\gg k_{\mathrm{B}}T$, the effective mass is given by the familiar expression, $m=\mu/v_0^2$. In Sec.7.2.2, we will use the approach outlined above to describe momentum drag.

The two-fluid model can also be used to describe energy transport. Indeed, particle

flow is accompanied by heat flow, described by

$$\mathbf{j}_{\rm q} = TS_e n_e \mathbf{V}_e + TS_h n_h \mathbf{V}_h \tag{7.21}$$

where S_e and S_h is the entropy per carrier for electrons and holes. Here the entropy and particle density can be related to the distribution function via

$$\begin{aligned} S_i &= -\frac{4k_{\rm B}}{n_i}\int \frac{d^2p}{(2\pi)^2}\left[(1-f_i(\mathbf{p}))\ln(1-f_i(\mathbf{p}))\right. \\ &\quad \left. +f_i(\mathbf{p})\ln f_i(\mathbf{p})\right], \quad n_i = 4\int \frac{d^2p}{(2\pi)^2} f_i(\mathbf{p}). \end{aligned} \tag{7.22}$$

In our analysis, we will need the value at charge neutrality. Direct numerical integration in Eq.(7.22) gives $S \approx 3.288 k_{\rm B}$. In Sec.7.2.3, we will connect $\mathbf{j}_{\rm q}$ to electric current, which will lead to a simple model for drag originating from E-mechanism.

7.2.2 Momentum drag mechanism

Here we will use the two-fluid model introduced in Sec.7.2.1 to derive momentum drag. To facilitate the analysis of transport equations, it is convenient to switch from vector notation to a more concise complex-variable notation. We will describe velocity, momentum and electric field by complex variables,

$$\tilde{V} = V_x + iV_y, \quad \tilde{P} = P_x + iP_y, \quad \tilde{E} = E_x + iE_y. \tag{7.23}$$

The solution of Eq. (7.19) can be written in a compact form by introducing the complex-valued quantities

$$\lambda_i = \frac{n_i}{\frac{m_i}{\tau_i} - iq_iB + \eta N}, \quad N = \sum_{i'=1\ldots4} n_{i'}. \tag{7.24}$$

Solving the transport equations [3] and summing electron and hole contributions to the electric current in each layer we obtain the current-field relation for the two layers using

[3]This algebraic exercise is made a lot easier by using the λ_i above. We can now express the velocity for

a 2×2 matrix that couples variables in layer 1 and layer 2:

$$\begin{pmatrix} \tilde{j}_1 \\ \tilde{j}_2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \tilde{E}_1 \\ \tilde{E}_2 \end{pmatrix}. \tag{7.25}$$

Here $\sigma_{11} = e^2\left[(\lambda_{1e} - \lambda_{1h})f_1 + \lambda_{1e} + \lambda_{1h}\right]$, $\sigma_{12} = \sigma_{21} = e^2(\lambda_{1e} - \lambda_{1h})f_2$,

$\sigma_{22} = e^2\left[(\lambda_{2e} - \lambda_{2h})f_2 + \lambda_{2e} + \lambda_{2h}\right]$,

$$f_1 = \frac{\eta(\lambda_{1e} - \lambda_{1h})}{1 - \eta\Lambda}, \quad f_2 = \frac{\eta(\lambda_{2e} - \lambda_{2h})}{1 - \eta\Lambda}, \tag{7.26}$$

where $\Lambda = \sum_i \lambda_i$. Here the quantities σ_{11} and σ_{22} describe the conductivity of layers 1 and 2, whereas the quantities σ_{12} and σ_{21} describe mutual drag between the layers (we note that $\sigma_{12} = \sigma_{21}$). The real and imaginary parts of σ_{12} describe the longitudinal and Hall drag.

The longitudinal and transverse drag resistivity can be obtained by inverting the matrix, Eq.(7.25), giving

$$\rho_{\rm d} = \rho_{\rm d,\parallel} + i\rho_{\rm d,Hall} = -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}}. \tag{7.27}$$

The quantities $\rho_{\rm d,\parallel}$ and $\rho_{\rm d,Hall}$ give the magnetodrag and Hall drag shown in Figs.7.4,7.5. This quantity features an interesting dependence on carrier density and magnetic field. We

---

each component as

$$\tilde{V}_i = \frac{\lambda_i}{n_i}\left[F + q_i\tilde{E}_i\right], \quad F = \eta\sum_i n_i\tilde{V}_i.$$

Combining coupled equations for $\tilde{V}_i$, yields a relation

$$\left[1 - \sum_{i'}\eta\lambda_{i'}\right]\left(\sum_i n_i\tilde{V}_i\right) = \sum_i q_i\lambda_i\tilde{E}_i$$

The quantity F can be found from above as

$$F = \frac{\eta}{1 - \eta\Lambda}\sum_i q_i\lambda_i\tilde{E}_i, \quad \Lambda = \sum_{i'=1\ldots4}\lambda_{i'}$$

Using this result, we evaluate electric current in each layer by summing the electron and hole contributions. For example, for layer 1 we have

$$\tilde{j}_1 = en_{1e}\tilde{V}_{1e} - en_{1h}\tilde{V}_{1h} = e\lambda_{1e}(F + e\tilde{E}_1) - e\lambda_{1h}(F - e\tilde{E}_1).$$

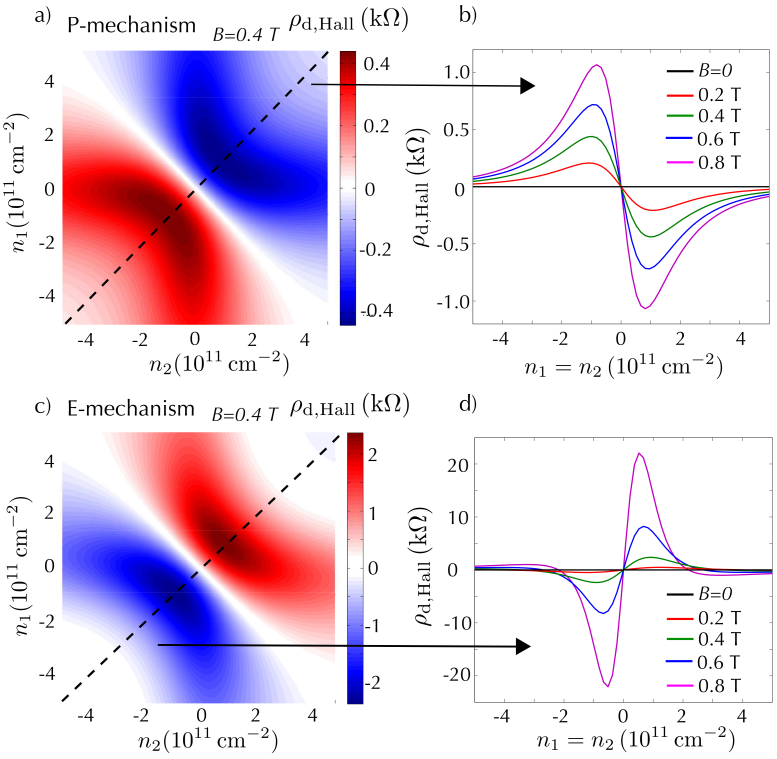

Figure 7.5: Hall drag originating from P and E mechanisms for the same parameters as in Fig.7.4. (a) Hall drag density dependence for P-mechanism, obtained from Eq.(7.27). (b) Line traces along black dashed line in (a) for various values of B. (c) Hall drag density dependence for E-mechanism, obtained from Eq.(7.34). (d) Line traces along black dashed line in (c) for various values of B. The difference in sign for the two contributions makes them easy to identify experimentally. Parameters used identical to Fig. 7.4.

will analyze the limit of small η (weak interactions). In this case, we have $\lambda_i \approx n_i/(\frac{m_i}{\tau_i} - iq_iB)$. This gives the drag resistance

$$\rho_{\rm d} \approx -\eta \frac{(\lambda_{1e}-\lambda_{1h})(\lambda_{2e}-\lambda_{2h})}{e^2(\lambda_{1e}+\lambda_{1h})(\lambda_{2e}+\lambda_{2h})}. \tag{7.28}$$

For $B = 0$, this quantity vanishes at DNP, $n_{1e} = n_{1h}$, $n_{2e} = n_{2h}$. Drag is negative for equal-polarity doping and positive for opposite-polarity doping, reproducing behavior well-known for momentum drag.

For nonzero B, the Hall drag and magnetodrag can be obtained by expanding $\mathrm{Im}\,\lambda_i(B) = \lambda_i(0)(q_i\tau_i/m_i)B + O(B^3)$ in Eq.(7.28). This gives Hall drag that vanishes exactly at DNP but is nonzero near DNP. For electron and hole densities near DNP, such that $\lambda_e \approx \lambda_h$, we find

$$\rho_{\rm d,Hall} = -\frac{\eta\tau}{em} B \left(\frac{(\lambda_{1e}-\lambda_{1h})}{(\lambda_{1e}+\lambda_{1h})} + \frac{(\lambda_{2e}-\lambda_{2h})}{(\lambda_{2e}+\lambda_{2h})} \right) + O(B^3). \tag{7.29}$$

This expression vanishes on the line $n_1 = -n_2$ corresponding to doping of opposite polarity in the two layers.

In contrast to Hall drag, magnetodrag is nonzero at DNP. From Eq.(7.28) we obtain a finite magnetodrag of a *positive sign:*

$$\rho_{\rm d,\parallel} = \eta \frac{\tau^2 B^2}{m^2} + O(B^4). \tag{7.30}$$

Here the quantities τ and m are evaluated at charge neutrality, $n_e = n_h$, in each layer. Interestingly, the magnetodrag sign comes out opposite to the sign predicted by the energy transport model (see below). The magnetodrag sign therefore provides a clear signature which discriminates between the E and P mechanisms in experiments.

The density dependence of magnetodrag and Hall drag predicted from P-mechanism is shown in Figs.7.4,7.5. In agreement with the above analysis, $\rho_{\rm d,\parallel}$ in Fig. 7.4a,b is positive

at DNP, increasing quadratically with B field. Also, $\rho_{\mathrm{d,Hall}}$ in Fig. 7.5a,b increases linearly with B field vanishing along $n_1 = -n_2$ as expected.

The plots were obtained by numerically evaluating the expression in Eq.(7.27), using parameter values described in Fig. 7.4 caption. In all our plots, we found it convenient to account for thermal and disorder broadening of the Dirac point in the same way by setting an effective temperature $T_{\mathrm{eff}} = T + \Delta$ in the evaluation of mass and entropy per carrier. We chose a disorder broadening $\Delta = 200\,\mathrm{K}$ that corresponds to a Dirac point width $\Delta n \approx 5 \times 10^{10}\mathrm{cm}^{-2}$ seen in the ultra-clean G/hBN/G devices used for drag measurements [44]. For simplicity, we also set the scattering rate at neutrality $\tau^{-1}(\mu = 0, T = 0) = \Delta/\hbar$ [see Section 7.2.4 for further discussion].

We parenthetically note that, while this model reproduces the qualitative features of P-mechanism, it is only valid not too far from DNP. In particular, we have ignored screening which becomes important far away from the Dirac point. As a result, P-drag seen in Figs. 7.4,7.5 does not diminish with doping. Accounting for screening of the interlayer interaction would generate suppression with doping, in agreement with previous studies of P-drag. [135–140]

7.2.3 Energy-driven drag mechanism

Here we analyze the contribution to drag resulting from energy transport (E-drag). We will start with evaluating the heat current $\mathbf{j}_{\mathrm{q}}$ [Eq.(7.21)] transported by electric current. In doing so, it will be instructive to first ignore the mutual drag effect discussed above, setting $\eta = 0$, and restore finite η later. Continuing to use complex variables for velocities and fields, we find $\tilde{V}_e = (\lambda_e q_e/n_e)\tilde{E}$, $\tilde{V}_h = (\lambda_h q_h/n_h)\tilde{E}$. Combining with Eq.(7.21), we find a relation

$$\tilde{j}_{\mathrm{q}} = Q\tilde{j}, \quad Q = \frac{T\left(S_e\lambda_e q_e + S_h\lambda_h q_h\right)}{\lambda_e q_e^2 + \lambda_h q_h^2}, \tag{7.31}$$

where $S_{e(h)}$ can be evaluated using the expression in Eq. 7.22. This relation is particularly useful since the effect of the Lorentz force is fully accounted for via λ_i. The imaginary part of Q describes the angle between the angle between the heat current and electric current, $\tilde{j}_{\rm q}$ and $\tilde{j}$, which corresponds to the Nernst/Ettingshausen effect.

Energy transport, described by Eqs.(7.21),(7.31), creates a temperature gradient across the system. For two layers in close proximity, fast heat transfer between layers due to interlayer electron scattering leads to a temperature profile which is essentially identical in both layers. [162] The temperature gradients can drive a local thermopower via

$$E = \frac{Q}{T}\nabla T, \tag{7.32}$$

where the quantity Q is given by the ratio of the heat current and electric current for the layer in question. As discussed in detail in Ref. [175] this relation follows from Onsager reciprocity combined with Eq.(7.31).

The temperature gradient can be found from balancing the heat flux due to thermal conductivity against the net heat flux due to electric current in the two layers, $j_{\rm q} = j_{1,{\rm q}} + j_{2,{\rm q}}$. While the details of the analysis somewhat depend on sample geometry (see Ref. [175] and discussion below), here we adopt a simplistic viewpoint and write the balance condition in a general algebraic form as

$$(\kappa_1 + \kappa_2)\nabla T = \beta j_{\rm q}, \tag{7.33}$$

where ∇T is the average temperature gradient across the system, κ_1 and κ_2 are thermal conductivities of the layers. The quantity $0 < \beta \leq 1$ is introduced to account for the "active part" of the heat that is not lost to contacts and/or the crystal lattice.

We will first discuss the general behavior that can be understood directly from Eq.(7.33) without specifying β value. Combining Eq.(7.33) and Eq.(7.32) to evaluate drag

voltage, we can write drag resistivity as

$$\rho_{12} = \beta \frac{Q_1 Q_2}{T(\kappa_1 + \kappa_2)}. \tag{7.34}$$

This quantity is symmetric under interchanging layers, $1 \leftrightarrow 2$. The real and imaginary parts of ρ_{12} describe magnetodrag and Hall drag. These quantities feature interesting dependence on carrier density shown in Fig.7.4c,d and Fig.7.5 c,d [see Fig. 7.4 caption for parameter values]. Notably, the signs of magnetodrag and Hall drag obtained from E-mechanism are opposite to those obtained from P-mechanism. The relation between the signs of the E and P contributions provides a convenient way to differentiate between the two mechanisms in experiment.

In our numerical simulations of E-mechanism we used $\eta = 0$ to reflect the "universal values" of drag that E-mechanism takes on in the weak coupling regime. For finite but small η, E-mechanism remains unaffected. We note, however, that at strong coupling, large values of η can affect the magnitude of E-mechanism drag.

The behavior of drag, described by Eq.(7.34), is particularly simple at charge neutrality. In this case, since $n_e = n_h$, $S_e = S_h$, the particle and hole contributions to the heat current $j_{\rm q}$ are of equal magnitude. Also, since $\lambda_e = \lambda_h$ at $B = 0$, drag resistivity vanishes at zero magnetic field B. Furthermore, at finite B the quantities λ_e and λ_h acquire a relative phase difference, such that $\lambda_e = \lambda_h^*$. As a result, the quantities Q_1 and Q_2 that enter Eq.(7.34) are purely imaginary, producing drag resistivity that has a negative sign for nonzero B. We can obtain magnetodrag by expanding in small B, which gives

$$\rho_{\rm d,\parallel} = -\beta \frac{TS^2}{2\kappa} \left(\frac{\tau}{m} B\right)^2, \tag{7.35}$$

where τ, m and κ are evaluated at charge neutrality. For an estimate, we will relate thermal conductivity to electrical conductivity using the Wiedemann-Franz relation, $\kappa = \frac{\pi^2 k_{\rm B}^2 T}{3e^2}\sigma$.

This relation is valid for degenerate Fermi systems, however we expect it to be also approximately valid near charge neutrality. This gives

$$\rho_{\mathrm{d},\parallel} \approx -\beta \frac{3S^2e^2}{2\pi^2\sigma k_{\mathrm{B}}^2}\left(\frac{\tau}{m}B\right)^2. \tag{7.36}$$

Comparing to the answer for P-mechanism, we find the ratio of the contributions due to momentum and energy mechanisms

$$\frac{\rho_{\mathrm{d},\parallel}^{(\mathrm{P})}}{\rho_{\mathrm{d},\parallel}^{(\mathrm{E})}} = -\frac{2\pi^2 k_{\mathrm{B}}^2}{3\beta S^2}\frac{\sigma\eta}{e^2}. \tag{7.37}$$

We can estimate entropy per carrier at DP by evaluating the integral over energy in Eq.(7.22). Using the value $S \approx 3.288 k_{\mathrm{B}}$ quoted above, we arrive at Eq.(7.16). Given the conductivity value at charge neutrality, $\sigma \approx 4e^2/h$, and taking into account that the mutual drag coefficient η is small when the fine structure constant $\alpha = e^2/\hbar v\kappa$ is small, [157, 158] $\eta \sim \alpha^2$, we conclude that the ratio in Eq.(7.16) is smaller than unity. This indicates that under very general conditions the E contribution overwhelms the P contribution in the DNP region.

The value of β in Eq.(7.33) depends on the rate of heat loss from electrons to the lattice and contacts. As an illustration, we consider the case when heat loss is dominated by cooling to the lattice. In this case, β depends on the relation between electron-lattice cooling length and system dimensions. We can model heat transport across the system as

$$\kappa(-\nabla^2 + \xi^{-2})T = -\nabla j_{\mathrm{q}}, \quad 0 < x < W, \tag{7.38}$$

where W is system width and ξ is the electron-lattice cooling length. Spatially uniform heat current j_{q} translates into a pair of delta-function sources, localized at $x = 0$ and $x = W$. Solving for the temperature profile, we obtain the temperature imbalance sustained between the sample edges, $\Delta T = T_{x=W} - T_{x=0} = \beta W j_{\mathrm{q}}/\kappa$, with the β value given by

$$\beta = \frac{1}{c}\mathrm{tanh}c, \quad c = \frac{W}{2\xi}. \tag{7.39}$$

This gives $\beta \to 1$ when $\xi \gg W$ (slow electron-lattice cooling) and $\beta \to 0$ when $\xi \ll W$ (fast electron-lattice cooling). The cooling length, ξ, in graphene can be as large as several microns for a wide range of temperatures up to room temperature. [26–28] For such temperatures, since in typical devices W is a few microns or smaller, the factor β can be close to unity, $\beta \sim 1$. However, the electron-lattice cooling rate grows at temperatures exceeding a few hundred kelvin owing to cooling pathway mediated by optical phonons. At such high temperatures, since the cooling length ξ shortens rapidly, Eq.(7.39) predicts vanishingly small β. In this case, temperature gradients in the electron system do not build up, rendering the E-mechanism ineffective. The latter regime (fast cooling) is not relevant, however, for practically interesting temperatures $T \lesssim 300\,\mathrm{K}$, where we expect $\beta \sim 1$ for few-micron-size devices.

We also note that β may be altered in a nontrivial way by boundary conditions, for example by contacts that act as heat sinks. In particular, in anisotropic systems or in systems with anisotropic contact placement, the relation between heat flow and ∇T can become anisotropic. In this case, β can be described as a 2×2 tensor (see Sec. 7.1 and Ref. [163] supplementary information for further discussion). While the qualitative behavior discussed above (drag order of magnitude and sign at DNP) is not expected to be altered by anisotropy in heat loss, the tensor character of β can affect the layer symmetry of the resultant drag [e.g. see Sec. 7.1 where the lack of symmetry $n_1 \leftrightarrow n_2$ stems from anisotropic device geometry]. In contrast, in the isotropic case, where heat flow is not influenced by device geometry or contact placement, β is a c-number. In this case, Eq.(7.34) predicts drag obeying layer symmetry, $n_1 \leftrightarrow n_2$.

Finally, we comment on the anomalously large values of $\rho_{\mathrm{d},\parallel}$ at the highest B fields seen in Fig. 7.4(d). These values far exceed P-mechanism, however they also exceed the in-plane sheet resistivity. This signals that our treatment, while successfully capturing E-

mechanism drag for low B, ceases to be valid for higher B. We note in this regard that for $B = 0.8\,\mathrm{T}$ the energy of the first Landau level, $E_1 \approx 380\,\mathrm{K}$, exceeds our disorder broadening value $\Delta = 200\,\mathrm{K}$. This hints at the importance of Quantum Hall physics at such fields.

7.2.4 Modeling procedure

Here we comment on the quantities that enter the two-fluid description, and discuss the sensitivity of the results to the simplifying assumptions made in the model.

In the two-fluid model we describe the momentum-velocity relation for each component as $\mathbf{P} = m_i \mathbf{V}$, where m_i is an "effective mass." An explicit expression for m_i as a function of T, μ can be found by expanding the distribution functions (Eq. 7.18) to lowest non-vanishing order in $\mathbf{a}_{e(h)}$:

$$m_i = \frac{1}{v_0} \frac{\int d^2\mathbf{p}\, p_x \nabla_{\mathbf{a}_x} f_i(\mathbf{p})}{\int d^2\mathbf{p}\, \frac{p_x}{p} \nabla_{\mathbf{a}_x} f_i(\mathbf{p})} = \frac{1}{v_0} \frac{\int d^2\mathbf{p}\, p_x^2 g_i(\mathbf{p})}{\int d^2\mathbf{p}\, \frac{p_x^2}{p} g_i(\mathbf{p})}, \tag{7.40}$$

where $g_i(\mathbf{p}) = f_i(\mathbf{p})(1 - f_i(\mathbf{p}))$.

The times τ_i for disorder scattering and carrier densities n_i in Eq. 7.19 are expressed through the distribution function (Eq. 7.18) with $\mathbf{a}_i = 0$:

$$\frac{1}{\tau_i} = \frac{4}{n_i} \int \frac{d^2\mathbf{p}}{(2\pi)^2} \frac{f_i(\mathbf{p})}{\tau_i(\epsilon_{\mathbf{p}})}, \quad n_i = 4 \int \frac{d^2\mathbf{p}}{(2\pi)^2} f_i(\mathbf{p}), \tag{7.41}$$

where the factor of four accounts for spin-valley degeneracy in each layer. We pick a model for the transport scattering time $\tau(\epsilon)$ to account for the experimentally observed linear dependence of conductivity vs. doping, $\sigma = \mu_* |n|$, where μ_* is the mobility away from the DP. This is the case for Coulomb impurities or strong point-like defects, such as adatoms or vacancies [178]. In both cases the scattering time has an approximately linear dependence on particle energy,

$$\tau(\epsilon)_{|\epsilon| \gtrsim \gamma} = \hbar |\epsilon| / \gamma^2, \quad \gamma = v_0 \sqrt{e\hbar/\mu_*} \tag{7.42}$$

where the disorder strength parameter γ is expressed through mobility. The value $\mu_* = 6 \cdot 10^4\,\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ measured in graphene on BN [10] yields $\gamma \approx 120\,\mathrm{K}$. Similar values for γ are obtained from the DP width extracted from the resistivity density dependence, [44] $\Delta n \approx 5 \cdot 10^{10}\,\mathrm{cm}^{-2}$.

In doing simulations, we found it convenient to use a different, simplified model for transport scattering which does not involve integration over particle distribution, yet yields results similar to those obtained from a more microscopic model, Eq.(7.41). We model the scattering time in the full range of doping densities as

$$\tau_i = (m_i v_0^2 + \Delta)\frac{\hbar}{\Delta^2} \tag{7.43}$$

where m_i depends on temperature and density via Eq.(7.40), and the parameter Δ describes the smearing of DP due to disorder. This model accounts for the experimentally observed linear dependence of conductivity vs. doping. In the simulation we used the value $\Delta = 300\,\mathrm{K}$ which translates into DP width of about $\Delta n \approx 10^{11}\,\mathrm{cm}^{-2}$, consistent with the above estimates.

Additionally, we found it convenient to account for disorder broadening of DP by using an effective temperature $T_{\mathrm{eff}} = T + \Delta$ in the evaluation of the effective mass in Eq.(7.40) and the entropy per particle in Eq. 7.22. This simple procedure captures the essential characteristics of DP broadening since smearing of the density of states by temperature and disorder occur in a similar fashion.

Using the parameters $\beta = 1$, $\eta = 0.23\hbar$, $\Delta = 300\,\mathrm{K}$, $T = 200\,\mathrm{K}$ and the model for scattering time τ_i in Eq.(7.43), we plot $\rho_{\mathrm{d},\parallel}$ and $\rho_{\mathrm{d,Hall}}$ for P-mechanism [Eq. 7.27] and E-mechanism [Eq.7.34] in Figs. 7.4, 7.5 respectively. As discussed above, this gives density dependence of drag that differs in sign for the two mechanisms. Additionally, we find that the E-mechanism magnitude exceeds that of P-mechanism for the region near DNP. This

agrees with the small ratio for P-mechanism vs. E-mechanism derived for small η.

7.3 Summary

In summary, we argue that magneto and Hall drag in graphene near charge neutrality is dominated by energy transport effects (E-mechanism) arising due to fast interlayer energy relaxation that couples to lateral energy flow and, via thermopower, drives electric current. This provides a description of the fate of Energy-driven drag in the presence of a magnetic field (introduced in the previous chapter for $B = 0$); E-drag continues to dominate at CN, changing sign and quickly attaining large values as B field is increased.

To compare E- and P- mechanisms on an equal footing, we developed a two-fluid framework which accounts both for the E-mechanism as well as for the standard momentum-transfer drag (P-mechanism), capturing the essential features of the two mechanisms. This unified approach is particularly instructive, not only because it produces both P-mechanism and E-mechanism, but also because it allows an unbiased way of comparing the magnitudes of the two mechanisms. Strikingly, the P and E mechanisms yield opposite sign for both magnetodrag and Hall drag resistivities. Along with a strong peak in magnetodrag at DNP originating from E-mechanism, this sign difference provides a clear way to experimentally distinguish the two mechanisms. Furthermore, the sensitivity of E-mechanism to boundary conditions, such as the placement of contacts that act as heat sinks, allows a new way of probing the two mechanisms - by checking the layer symmetry (cf. Eq. 7.14).

We show that the magnitude of drag originating from the two mechanisms is dominated by very different effects. The P mechanism is mostly controlled by the interlayer electron-electron interaction, becoming weak when this interaction decreases due to large layer separation or screening. In contrast, the E mechanism is controlled by long-range

energy transport, yielding a "universal value" for drag: the E contribution is essentially independent of the interlayer carrier scattering rate so long as it is faster than electron-lattice cooling. Slow electron-lattice cooling in graphene ensures that drag near DNP can remain large even when interlayer electron interactions are weak. This makes graphene an ideal system to observe E-drag and thereby probe energy transport on the nanoscale.

Chapter 8

Energy waves and Ballistic Heat Transfer in Graphene

In the previous chapters, we examined how hot carriers give rise to new charge transport behavior. In this chapter[1], we examine novel energy transport at charge neutrality. In particular, we ask: Can heat in graphene propagate by directed ballistic pulses? While energy transport is typically of a diffusive character, some materials can display a very different heat transfer mode — energy pulses transmitted in a collective wave-like fashion. For such pulses the distance travelled scales linearly with the travel time, $\Delta x = v\Delta t$. Thermal waves, called second sound, can occur in solids that host a "thermal liquid" of phonons. [179] The existence of second sound requires that its frequency satisfies $\gamma_p \ll \omega \ll \gamma_N$, where γ_N and γ_p are the phonon momentum-conserving (normal) and momentum-nonconserving (Umklapp) scattering rates. The rate γ_N grows with temperature, however the rate γ_p, grows even faster. For this reason second sound, originally discovered in superfluid He, [180] was observed only in a handful of solids, namely solid He, NaF and Bi. [181] Second sound speed is close to $s' = s/\sqrt{3}$, where s is the sound velocity, giving values such as 20 m/s for ^{4}He and 780 m/s for Bi. The relatively low velocity values facilitate experimental detection of

[1]This work was carried out with Trung Van Phan and Leonid Levitov. It has been posted on the arXiv preprint server as TV Phan, JCW Song, LS Levitov, arXiv: 1306:4972 (2013).

second sound, yet they also limit its utility in energy transduction applications.

Can heat transfer occur at supersonic speeds? So far supersonic heat transfer has not been known in an earthly setting. However, theory predicts energy and entropy waves in interacting systems of relativistic particles. [182] These long-wavelength oscillations, sometimes called *cosmic sound*, propagate with very high velocity $c' = c/\sqrt{3}$, where c is the speed of light (the quantity under square root is the dimensionality of space). Acoustic oscillations obeying this relation underpin the modern interpretation of Cosmic Microwave Background Radiation, a relict of the "big bang" creation of the universe [183, 184]. The relation $\Delta x = c'\Delta t$ serves as a "standard ruler" for relating cosmic length scales and distant times.

8.1 Energy waves in Graphene at Charge Neutrality

In this chapter, we predict waves analogous to cosmic sound in graphene. Electrons in graphene behave as relativistic particles moving with velocity $v \approx 10^6\,\mathrm{m/s}$. At charge neutrality graphene hosts charge-compensated plasma with strong interactions. The carrier-carrier scattering in this system, under typical conditions, is much faster than electron-lattice cooling rate and momentum relaxation due to disorder scattering, which makes graphene ideal for realizing an electronic analog of second sound. Rapid exchange of energy and momentum among colliding particles results in energy propagating as a collective excitation shared by many particles. Furthermore, since momentum is conserved alongside with energy, energy transport is characterized by inertia (finite momentum associated with energy flux). As we will see, coupled energy and momentum transport gives rise to wave-like propagation of energy with a characteristic velocity

$$v' = \frac{v}{\sqrt{2}} \approx 0.71 \cdot 10^6\,\mathrm{m/s}, \tag{8.1}$$

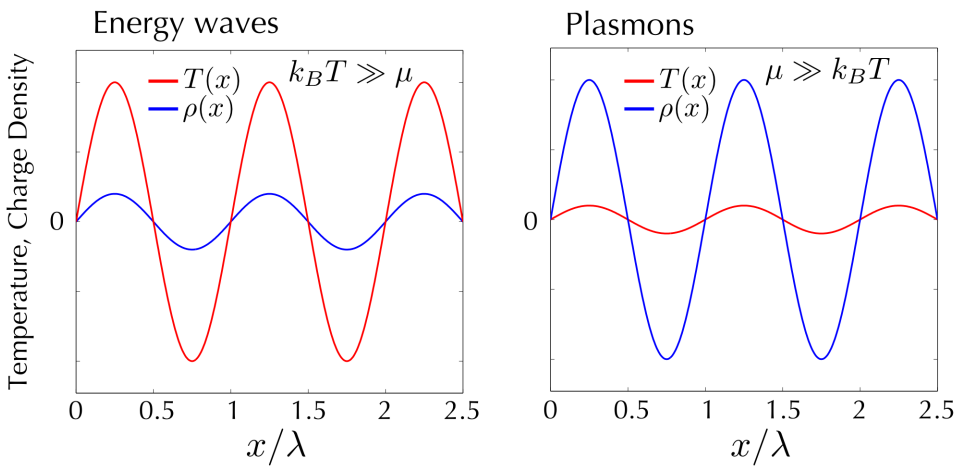

Figure 8.1: Oscillating temperature profile that results from a collective energy wave. Energy waves are decoupled from charge dynamics at charge neutrality, $\mu = 0$, become charge-coupled upon doping away from neutrality, $\mu \neq 0$, and eventually morph into plasmons at large μ. Coupling to plasmons, Eq.(8.19), can enable all-electric detection of energy waves.

where v is the Fermi velocity. The value v' is about 10^3 times larger than the velocity for the phonon mechanism.

Pulse-echo measurements and standing wave resonances were used to probe phonon second sound. [180, 181] For electronic second sound in graphene, gate-tunability of carrier density provides a new knob to probe ballistic energy transfer, having no analog in phonon second sound or cosmic sound. We will see that energy waves, which are fully uncharged at charge neutrality, become coupled to charge dynamics upon doping (see Fig. 8.1). Coupling to plasmons provides an all-electric way to excite and detect energy waves. One promising approach is the recently developed spatially resolved nanoscale probe of plasmonic standing waves. [185, 186] Local probes can be used to excite and detect energy waves using their coupling to plasmons near charge neutrality, $k_BT \gg \mu$. Alternatively, a pulse echo technique similar to that used for detecting ballistic electron resonances in carbon nanotubes [95] can be used.

The origin of collective thermal waves can be understood on very general grounds. We will first discuss cosmic sound, and then generalize to electronic waves in graphene. The

universal behavior of relativistic thermal waves follows from the proportionality between the energy flux and momentum density of a relativistic gas, $\mathbf{j}_\epsilon = c^2\mathbf{p}$, and the relativistic pressure-energy relation, $P = \frac{1}{3}W$. The wave equation follows under very general assumptions from energy and momentum conservation written in terms of the 4×4 stress tensor:

$$\partial_i T_{ij} = 0, \quad T = \begin{pmatrix} W & \mathbf{j}_\epsilon \\ \mathbf{j}_\epsilon & \mathbf{P} \end{pmatrix}, \tag{8.2}$$

where $\mathbf{P}$ is a shorthand for a diagonal 3×3 matrix $P\delta_{\alpha\beta}$, and $i = \{0, 1, 2, 3\}$, where 0 label time and $1, 2, 3$ denote space. Acoustic oscillations obeying the law $c' = c/\sqrt{3}$ arise directly from the coupled dynamics of energy and momentum governed by Eq.(8.2).

The origin of energy waves in graphene can be understood in direct analogy with cosmic sound. This is done most easily starting from conservation laws for energy and momentum of the electron system, described by the continuity equations

$$\partial_t W + \nabla \mathbf{j}_\mathrm{q} = -\gamma_w(W - W_0), \quad \partial_t p_i + \nabla_j \sigma_{ij} = -\gamma_p p_i. \tag{8.3}$$

Here $\mathbf{j}_\mathrm{q}$ is the heat current, and σ_{ij} is the stress tensor. The rates of momentum relaxation and electron-lattice cooling, γ_p and γ_w, are introduced to account for disorder scattering and energy loss to the lattice. For simplicity, we suppressed coupling to electric field, which limits the present discussion to charge neutrality where energy and momentum oscillations are decoupled from charge oscillations. This coupling will be reinstated below when we discuss the effect of doping, and the relation between thermal waves and plasmons.

As we will see, the quantities $\mathbf{j}_\mathrm{q}$ and σ_{ij} depend on momentum density and temperature in a way mimicing that in relativistic gas. This dependence generates a cross-coupling between the T and $\mathbf{p}$ dynamics, leading to oscillations and wave propagation. Specifically,

$$\mathbf{j}_\mathrm{q} = v^2\mathbf{p} - \kappa\nabla T, \quad \sigma_{ij} = \frac{1}{2}W\delta_{ij} - \nu_1\nabla_j p_i - \nu_2\delta_{ij}\nabla_k p_k, \tag{8.4}$$

were v is the velocity, $W(T) \sim T^3$ is the energy density, κ is thermal conductivity, and $\nu_{1,2}$ are viscosity parameters. The first terms in $\mathbf{j}_{\rm q}$ and σ_{ij} generate a nontrivial coupling of the T and $\mathbf{p}$ dynamics. Since these terms do not depend on gradients, they provide a dominant contribution in the long-wavelength limit: suppressing the terms in Eq.(8.4) first-order in gradients, we find coupled equations $\partial_t W + v^2 \partial_j p_j = 0$, $\partial_t p_i + \frac{1}{2}\partial_i W = 0$. This yields a wave equation $(\partial_t^2 - \frac{1}{2}v^2\nabla^2)W = 0$ describing energy propagation with the velocity given in Eq.(8.1). In the absence of damping, wave-like energy transport results in *ballistic* energy propagation, whereby distance travelled scales linearly with time, $\ell \sim v't$.

The relation between momentum density and energy flux which is key in the above analysis can be obtained by analyzing a homogeneous macroscopic flow of Dirac particles. This is described by a distribution

$$n(\mathbf{p}) = \frac{1}{e^{\beta(\epsilon_{\mathbf{p}}^{(\pm)} - \mathbf{up} - \mu)} + 1}, \quad \epsilon_{\mathbf{p}}^{(\pm)} = \pm v|\mathbf{p}|, \tag{8.5}$$

with $\mathbf{u}$ the velocity. At charge neutrality, $\mu = 0$, we evaluate momentum density at first order in $\mathbf{u}$. After some algebra we find

$$\langle \mathbf{p} \rangle = \sum_{\mathbf{p}} \mathbf{p}\delta n(\mathbf{p}) = N\frac{9\zeta(3)p_T^3}{4\pi\hbar^2 v}\mathbf{u}. \tag{8.6}$$

Here $\sum_{\mathbf{p}} ... = N\sum_{\pm}\int ...\frac{d^2p}{(2\pi\hbar)^2}$, $N = 4$ is the number of spin/valley species, and $p_T = k_{\rm B}T/v$. The sum is taken over both Dirac bands, $\epsilon_{\mathbf{p}}^{(\pm)} = \pm v|\mathbf{p}|$, and the quantity $\delta n(\mathbf{p}) = -(\mathbf{up})\partial f/\partial\epsilon$ is a first-order variation in $\mathbf{u}$. Next, evaluating the energy flux at first order in $\mathbf{u}$ and comparing to Eq.(8.6), we find

$$\mathbf{j}_{\rm q} = \sum_{\mathbf{p}} \epsilon_{\mathbf{p}}^{(\pm)} \mathbf{v}\delta n(\mathbf{p}) = N\frac{9\zeta(3)p_T^3 v}{4\pi\hbar^2}\mathbf{u} = v^2\langle \mathbf{p} \rangle \tag{8.7}$$

which determines the coupling of momentum to the heat flux, Eq.(8.4). This relation captures the essence of the coupling between collective particle motion and energy flow in the system.

8.2 Damping and Frequency Range

Various damping mechanisms can be most easily assessed by analyzing solutions for Eqs.(8.3). Taking the quantities $\mathbf{j}_{\rm q}$ and σ_{ij} in the form given in Eq.(8.4), and constructing plane-wave solutions, $W, \mathbf{p} \sim e^{-i\omega t + i\mathbf{kx}}$, we find the dispersion relation

$$(\omega + iD\mathbf{k}^2 + i\gamma_p)(\omega + i\nu\mathbf{k}^2 + i\gamma_w) = \frac{v^2}{2}\mathbf{k}^2 \tag{8.8}$$

where we defined diffusivity $D = \kappa/C$, where C is specific heat, and $\nu = \nu_1 + \nu_2$. Eq.(8.8) indicates that thermal waves are heavily damped at very high and very low frequencies. Damping at high frequencies is dominated by the diffusive and viscous terms, whereas damping at low frequencies is dominated by γ_p and γ_w [2].

To understand the relation between various timescales, we will use parameter values estimated for pristine graphene samples which are almost defect free, such free-standing graphene. [187] Disorder scattering can be estimated from the measured mean free path values which reach a few microns at large doping [188]. Using the momentum relaxation rate square-root dependence on doping, $\gamma_p \propto n^{-1/2}$, and extrapolating to charge neutrality, $n \sim 10^{10}\,\mathrm{cm}^{-2}$, gives values $\gamma_p^{-1} \sim 0.5\,\mathrm{ps}$. Theory predicts slow electron cooling in pristine graphene at neutrality, giving $\gamma_w^{-1} \sim T^{-2}$ with predicted values as large as 10 ns for $T = 100\,\mathrm{K}$ [11]. Cooling times can be shortened in the presence of disorder scattering [26], however, under any circumstances, cooling is expected to be much slower than momentum relaxation, $\gamma_w \ll \gamma_p$. The carrier-carrier scattering which dominates the hydrodynamical regime has little effect on cooling.

Refs. [157, 158] estimate the carrier-carrier scattering rate as $\gamma_N \approx A\alpha^2 k_{\rm B}T/\hbar$, where α is the interaction strength. For $T = 100\,\mathrm{K}$, approximating the prefactor as $A \approx 1$

[2] This can be easily seen by noting that energy waves have dispersion $\omega = v'|\mathbf{k}|$ and looking for the terms which dominate at high and low ω, being the $Dk^2, \nu k^2$, and γ_p, γ_w terms respectively.

[157, 158], we obtain characteristic times, $\gamma_N^{-1} \approx 70\,\mathrm{fs}$, which are much shorter than the values γ_p^{-1} and γ_w^{-1} estimated above. The inequalities $\gamma_w \ll \gamma_N$, $\gamma_p \ll \gamma_N$ justify our hydrodynamical description of transport. Estimating the diffusion constant and viscosity, D, $\nu \approx \frac{1}{2}v^2/\gamma_N$, and comparing to the dispersion relation, Eq.(8.8), we see that damping is weak in the range of frequencies defined as

$$\gamma_p \lesssim \omega \lesssim \gamma_N$$

Using the values estimated above, gives frequencies $f = \omega/2\pi$ in the low THz range, $0.1\,\mathrm{THz} \lesssim f \lesssim 3\,\mathrm{THz}$.

8.3 Energy Waves and Plasmons at Finite Doping

Next we will analyze the effect of finite doping. As we will see, the thermal waves morph into plasma waves upon doping away from neutrality. We will employ a microscopic approach based on the quantum kinetic equation, which will allow us to justify our hydrodynamical treatment by a microscopic analysis. The kinetic equation for electron distribution reads

$$(\partial_t + \mathbf{v}\nabla_{\mathbf{x}} + e\mathbf{E}\nabla_{\mathbf{p}})\, n(\mathbf{x}, \mathbf{p}, t) = I^N + I^{\mathrm{el-ph}} + I^{\mathrm{dis}}, \tag{8.9}$$

where $\mathbf{E}$ is the electric field, and I^N, $I^{\mathrm{el-ph}}$, I^{dis} describe (normal) two-particle collisions, electron-phonon collisions and scattering by disorder. The electric field can be either extrinsic (imposed externally) or intrinsic, arising due to long-wavelength charge fluctuations (as will be the case in our analysis of plasmons).

Collective transport arises in the regime dominated by normal collisions, when the processes $I^{\mathrm{el-ph}}$ and I^{dis} are negligible. The quantity I^N is given by the standard expression

which accounts for energy and momentum conservation in two-particle collisions, [160]

$$I^N[n] = \sum_{\mathbf{p}_1,\mathbf{p}',\mathbf{p}'_1} w_{\mathbf{p}',\mathbf{p}'_1\to\mathbf{p},\mathbf{p}_1}\left[(1-n(\mathbf{p}))(1-n(\mathbf{p}_1))\right.$$
$$\left.\times n(\mathbf{p}')n(\mathbf{p}'_1) - (1-n(\mathbf{p}'))(1-n(\mathbf{p}'_1))n(\mathbf{p})n(\mathbf{p}_1)\right]$$
$$\times\delta(\mathbf{p}+\mathbf{p}_1-\mathbf{p}'-\mathbf{p}'_1)\delta(\epsilon_{\mathbf{p}}+\epsilon_{\mathbf{p}_1}-\epsilon_{\mathbf{p}'}-\epsilon_{\mathbf{p}'_1}). \tag{8.10}$$

The transition rate w is given by the electron scattering vertex which includes coherence factors, e.g. Born approximation yields $w_{\mathbf{p}',\mathbf{p}'_1\to\mathbf{p},\mathbf{p}_1} = \frac{2\pi}{\hbar}|V_{\mathbf{p},\mathbf{p}_1,\mathbf{p}'\mathbf{p}'_1}|^2$.

In the long-wavelength limit we can analyze solutions of the kinetic equation, Eq.(8.9), perturbatively for a weak inhomogeneity. Setting the left-hand side of Eq.(8.9) to zero (electric field becomes small in the long-wavelength limit), we find that the kinetic equation is approximately solved by zero modes of the collision integral. The zero-mode equation $I^N[n] = 0$ can be solved by taking into account the energy-conserving and momentum-conserving character of the collision operator[3] A standard Boltzmannesque reasoning gives a general solution of the form [160]

$$n(\mathbf{x},\mathbf{p},t) = \frac{1}{e^{\beta(\mathbf{x},t)(\epsilon^{(\pm)}_{\mathbf{p}}-\mathbf{p}\cdot\mathbf{u}(\mathbf{x},t)-\mu(\mathbf{x},t)}+1}. \tag{8.11}$$

This equation determines the $\mathbf{p}$ dependence of particle distribution, leaving the dependence of the quantities $\mathbf{u}$, β and μ on position and time unspecified. Here $\mathbf{u}(\mathbf{x},t)$ is the hydrodynamical velocity describing collective motion of the e-h plasma, $\beta(\mathbf{x},t)$ describes temperature variation, and $\mu(\mathbf{x},t)$ describes the local chemical potential deviation from equilibrium. Accounting for variations in μ is inessential at charge neutrality, where energy and momentum oscillations occur at $\mu(\mathbf{x},t) = 0$. However, as we will see, μ must be included in the full analysis to achieve a unified description of both the undoped and doped regime. In particular,

[3]As a result of this choice - ie. only looking at the zero modes - this particular treatment does not capture viscous or diffusive terms described in 8.4; none of the terms below will depend on the electron-electron scattering rate. While these terms can be captured by a more general quantum kinetic equation treatment (see for eg. Ref. [189]), this is beyond the scope of this chapter.

plasma oscillations in the doped regime are associated with a time-varying μ.

The hydrodynamic description of energy and momentum transport can be obtained directly from conservation laws for particle density, energy and momentum. Specifically, we consider deviations in these quantities from equilibrium, $\delta N = \sum_{\mathbf{p}} \delta n(\mathbf{p})$, $\delta W = \sum_{\mathbf{p}} \epsilon_{\mathbf{p}}^{(\pm)} \delta n(\mathbf{p})$, $\langle \delta \mathbf{p} \rangle = \sum_{\mathbf{p}} \mathbf{p} \delta n(\mathbf{p})$. Integrating Eq.(8.9) over $\mathbf{p}$ yields the continuity equation for particle number. Multiplying Eq.(8.9) by $\epsilon_{\mathbf{p}}^{(\pm)}$ (or, by $\mathbf{p}$), integrating, and accounting for energy (momentum) conservation by normal collisions, yields equations for energy and momentum transport,

$$\partial_t \delta N + \nabla \mathbf{j}_N = 0, \quad \partial_t \delta W + \nabla \mathbf{j}_{\mathrm{q}} = 0, \tag{8.12}$$

$$\partial_t \langle \delta p_i \rangle + \nabla_j \sigma_{ij} - n_0 e E_i = 0. \tag{8.13}$$

Here $\mathbf{j}_N = \langle \mathbf{v} \rangle$ and $\mathbf{j}_{\mathrm{q}} = \langle \mathbf{v} \epsilon_{\mathbf{p}} \rangle$ are particle and energy currents, $\sigma_{ij} = \langle v_i p_j \rangle$ is the stress tensor, and n_0 is carrier density. In the interest of brevity we will denote averaging over the distribution variation δn by angular brackets, $\langle \mathcal{A} \rangle = \sum_{\mathbf{p}} \mathcal{A} \delta n(\mathbf{p})$, and suppress the $\pm$ superscript of $\epsilon_{\mathbf{p}}$. Here we used the continuity equation for δN to simplify the equation for δW by dropping μ: $\langle \epsilon_{\mathbf{p}} - \mu \rangle \to \langle \epsilon_{\mathbf{p}} \rangle$, $\langle \mathbf{v}(\epsilon_{\mathbf{p}} - \mu) \rangle \to \langle \mathbf{v} \epsilon_{\mathbf{p}} \rangle$.

The hydrodynamic equations at first order in deviations from equilibrium can be obtained by expanding the expression in Eq.(8.11) as

$$\delta n = \frac{\partial n}{\partial \beta} \delta\beta(\mathbf{x}, t) + \frac{\partial n}{\partial \mu} \delta\mu(\mathbf{x}, t) + \frac{\partial n}{\partial u_i} \delta u_i(\mathbf{x}, t), \tag{8.14}$$

plugging in the conservation laws and integrating over $\mathbf{p}$. In doing so it will be convenient to combine the first two terms, which are isotropic in $\mathbf{p}$, and denote them as $\delta_1 n$. The last term, which has angular dependence of the form $\mathbf{p} \cdot \mathbf{u}$, will be denoted as $\delta_u n$. Eq.(8.11) gives

$$\delta_1 n = \left(\frac{\epsilon - \mu}{\beta} \delta\beta - \delta\mu \right) \frac{\partial n}{\partial \epsilon}, \quad \delta_u n = -\mathbf{p} \cdot \delta\mathbf{u} \frac{\partial n}{\partial \epsilon} \tag{8.15}$$

Combining these expressions with the conservation laws, and taking into account the angular dependence in Eq.(8.15), we see that the quantities δN, δW and σ_{ij} are expressed through $\delta_1 n$, whereas $\mathbf{j}_N$, $\mathbf{j}_\mathrm{q}$ and $\langle \delta p_i \rangle$ are expressed through $\delta_u n$. Then, with angular averaging performed via $\langle \mathbf{v} \delta_u n \rangle = -\frac{1}{2}\mathbf{u}\langle \epsilon_\mathbf{p} \rangle$, the conservation laws for particle number and energy, Eq.(8.12), become

$$\partial_t \langle \delta_1 n \rangle = A \nabla \mathbf{u}, \quad \partial_t \langle \epsilon_\mathbf{p} \delta_1 n \rangle = B \nabla \mathbf{u}, \tag{8.16}$$

where $A = \langle \frac{\epsilon_\mathbf{p}}{2} \frac{\partial n}{\partial \epsilon} \rangle$, $B = \langle \frac{\epsilon_\mathbf{p}^2}{2} \frac{\partial n}{\partial \epsilon} \rangle$.

Lastly, we consider the momentum transport equation, Eq.(8.13). The stress tensor, after angular averaging performed as above, can be written as $\sigma_{ij} = \frac{1}{2}\delta_{ij}\langle \epsilon_\mathbf{p} \delta_1 n \rangle$. Expressing the electric field through spatial variation of particle density, we have $\mathbf{E} = -\nabla \int \frac{e}{|\mathbf{x}-\mathbf{x}'|} \langle \delta_1 n_{\mathbf{x}'} \rangle d^2 x'$. We note that σ_{ij} and $\mathbf{E}$ are expressed through the same quantities as those appearing under time derivatives in Eq.(8.16). With the help of this observation, we can write the time derivative of Eq.(8.13) as

$$\partial_t^2 \langle \delta \mathbf{p} \rangle + \frac{1}{2} B \nabla^2 \mathbf{u} = -n_0 \nabla \int \frac{e^2}{|\mathbf{x}-\mathbf{x}'|} A \nabla \mathbf{u}(\mathbf{x}') d^2 x'. \tag{8.17}$$

Finally, we express $\langle \delta \mathbf{p} \rangle$ through $\mathbf{u}$ by writing

$$\langle \delta \mathbf{p} \rangle = \langle \mathbf{p} \delta_u n \rangle = -\langle \frac{\mathbf{p}^2}{2} \frac{\partial n}{\partial \epsilon} \rangle \mathbf{u} = -\frac{B}{v^2}\mathbf{u}. \tag{8.18}$$

Plugging this in Eq.(8.17), and passing to Fourier harmonics, $\mathbf{u}(\mathbf{x}, t) \sim e^{i\mathbf{k}\mathbf{x} - i\omega t}$, we obtain the dispersion relation describing charge-coupled thermal waves:

$$\omega^2 = \frac{v^2}{2}\mathbf{k}^2 + 2\pi e^2 \lambda n_0 v^2 |\mathbf{k}|, \quad \lambda = \frac{A}{B}. \tag{8.19}$$

The plasmonic correction (second term) vanishes at charge neutrality, $\mu = 0$, and is small near it. Numerical analysis of Eq.(8.19) can be done using the integral $\int_0^\infty \frac{t^{s-1}}{e^t/z+1} dt = -\Gamma(s)\mathrm{Li}_s(-z)$, where Li_s is the polylogarithm function and $z = e^{\beta\mu}$. We find

$$\lambda n_0 = -\alpha \frac{(\mathrm{Li}_2(-z) - \mathrm{Li}_2(-1/z))^2}{\mathrm{Li}_3(-z) + \mathrm{Li}_3(-1/z)} \approx \frac{16(\ln 2)^2 \beta \mu^2}{9\pi\zeta(3)\hbar^2 v^2}, \tag{8.20}$$

with the last equation valid near charge neutrality $|\mu| \ll k_{\mathrm{B}}T$ [here $\alpha = \frac{2}{3\pi\hbar^2 v^2 \beta}$]. The last term in Eq.(8.19) becomes dominant far from neutrality, $|\mu| \gg k_{\mathrm{B}}T$. In this limit, we reproduce the standard (hydrodynamical) plasmon dispersion [111]. Charge coupling that is nonzero but small near charge neutrality, Eq.(8.19), provides a convenient tool for an all-electric excitation and detection of energy waves. For instance, using the recently scanning techniques developed in Ref. [185,186], the dispersion in Eq. 8.19 can be mapped; Eq. 8.19 predicts that deviations from the standard plasmon dispersion caused by energy waves are maximized close to charge neutrality.

In summary, graphene can host new collective modes, namely long-wavelength energy oscillations that propagate as weakly damped waves. Energy waves exist in the hydrodynamical frequency range, $\omega < \gamma_N$. This condition sets such waves apart from various types of collisionless collective modes proposed at charge neutrality, which occur at $\omega \geq \gamma_N$. [190–192] The electronic nature of energy waves ensures their high propagation velocity, which can be 10^3 times larger than the highest values known for the phonon mechanism. Directed ballistic energy pulses enabled by electronic second sound open the door to achieving high-speed energy transduction in solids.

Chapter 9

Afterword

In summary, we have shown how hot carriers can significantly alter the behavior of graphene and graphene heterostructures in a variety of different settings. This stems from slow electron-lattice and strong thermoelectricity in graphene, particularly close to Charge Neutrality.

In Chapter 1-5 we discussed hot carriers in the context of graphene's photoresponse. First in Chapter 2, we showed through a theoretical prediction and subsequent experiment (and analysis) that hot carriers mediate the photocurrent response in graphene, so that the primary way photocurrent is generated at a graphene p-n junction is through a Photothermoelecric effect - light heats the electrons at the junction which then subsequently drive a thermoelectric current. In Chapter 3, we detailed how hot carriers in graphene cool down to the lattice temperature. In particular, we showed how a new mechanism of hot carrier cooling called supercollisions - which relaxes momentum conservation - is responsible for hot carrier cooling over a wide range of technologically relevant temperatures including room temperature. In Chapter 4, we showed how hot carriers can be generated through photoexcitation and subsequent intraband carrier-carrier scattering (called Impact Excitation). We showed that the fast Impact Excitation process wins over optical phonon emission

to allow the ambient electrons in graphene to absorb a large fraction of the energy of the initial photoexcited electron, allowing graphene to be efficient at converting light energy to electronic heat. In Chapter 5, we laid out a framework for linking local photoexcitation to global photocurrent measured at far away contacts which mirrored the Shockley-Ramo theorem for semiconductors and vacuum tubes.

In Chapter 6 & 7, we showed how hot carriers can affect the transport properties of double layer graphene heterostructures. In particular, we showed how coulomb mediated vertical energy transfer between two layers can give rise to a new energy-driven drag mechanism. Interestingly, energy-driven drag gives rise to drag resistance that is insensitive to the strength of the Coulomb interaction in large enough devices. This new mechanism explains the anomalous peak in drag resistance at charge neutrality in recent experiments on double layer graphene heterostructures both at $B = 0$ and at $B \neq 0$.

Lastly in Chapter 8, we examined how a new ballistic electronic heat transport mode can arise near charge neutrality that manifests in wavelike energy waves. This occurs when graphene is in the hydrodynamic regime in which carrier-carrier scattering is far faster than momentum or energy relaxation, a property of hot carriers in graphene.

9.1 Outlook

While it has been ten years since graphene was discovered, graphene still continues to offer surprises and opportunities to find new physics. In particular, the recently established ability to stack various van der Waals' materials together [3, 10] may open up a new vista for exploring and engineering new electronic properties. However in this outlook, we restrict ourselves to discussing the horizons of hot carriers in graphene. Here we lay out some possible future directions.

One particular area that has yet to be fully understood is the photoexcitation cascade of *undoped* graphene[1]. Recently, carrier-multiplication events have stirred up scrutiny and some controversy in the field with various experimental groups claiming carrier multiplication events [193] and others that claim they see no evidence for it [103]. In Chapter 4, we discussed how kinematical constraints limit the type of processes that can relax a high energy photoexcited carrier. In particular for two-body collisions, simultaneous enegry and momentum conservation prevented interband (so-called Impact Ionization) processes from occurring. However, higher-order processes can alleviate this bottleneck and have yet to be fully explored. For example, a disorder-mediated process in the same spirit as supercollisions in Chapter 3 may allow for interband processes to relax high energy photoexcited carriers; numerical studies have indicated that higher fluence pumping of graphene can allow for carrier multiplication events to occur [57, 58]. How these happen and how they change the interaction induced lifetime of high energy particles remains unclear; a deeper theoretical understanding will give a fuller picture of the photoexcitation cascade in graphene.

Given the intense interest in making graphene bolometers [22,23], another challenge in the field is to more efficiently extract the energy from hot carriers. One strategy is to directly extract hot carriers in a vertical sandwich structure. Such a structure may allow for efficiencies to exceed to Shockley-Quiesser limit [2] and may be a realization of a hot carrier solar cell or an efficient photodetector. Further, since this would rely on extracting charge carriers faster than electron-lattice cooling processes, a new transport regime may be accessed - cooling by extracting carriers - and opens the door for new behavior in layered graphene structures under light illumination.

A further goal is to understand what happens to energy-driven drag in the quantum Hall regime. While some of the predictions of energy-driven drag have been confirmed

[1] In Chapter 4, we treated *doped* graphene

in experiment [44], these have focussed on the classically weak B field regime. Recent experiments [171] have revealed complicated structures at high B which at this time cannot be accounted for by the results presented in this thesis, and leaves the theoretical understanding of drag graphene heterostructures in high field wide open.

9.2 Last Words

The work undertaken in this thesis has been tremendously rewarding for me personally. The subfield of "Hot Carriers in Graphene" took off just as I started in it and has come of its own; some of the ideas described in this thesis have been adopted by workers in the field and I am very excited and gratified to see how far it has come. The close collaboration of theorists with experimentalists has been a hallmark of my time here, and I feel very privileged to have played some small part in advancing our common boundaries of knowledge.

Bibliography

[1] Pop, E. Energy dissipation and transport in nanoscale devices. *Nano Research* **3**, 147 (2010).

[2] Shockley, W. & Queisser, H. J. Detailed balance limit of efficiency of p-n junction solar cells. *Journal of Applied Physics* **32**, 510 (1961).

[3] Geim, A. K. & Grigorieva, I. V. Van der waals heterostructures. *Nature* **499**, 419 (2013).

[4] Bernardi, M., Palummo, M. & Grossman, J. Extraordinary sunlight absorption and one nanometer thick photovoltaics using two-dimensional monolayer materials. *NanoLetters* **13**, 3664 (2013).

[5] Bonaccorso, F., Sun, Z., Hasan, T. & Ferrari, A. C. Graphene photonics and opto-electronics. *Nature Photonics* **4**, 611 (2010).

[6] Nair, R. R. *et al.* Fine structure constant defines visual transparency of graphene. *Science* **320**, 1308 (2008).

[7] Xia, F., Müller, T., Lin, Y.-M., Valdes-Garcia, A. & Avouris, P. Ultrafast graphene photodetector. *Nature Nanotechnology* **4**, 839 (2009).

[8] Mueller, T., Xia, F. & Avouris, P. Graphene photodetectors for high-speed optical communications. *Nature Photonics* **4**, 297 (2010).

[9] Novoselov, K. S. *et al.* Electric field effect in atomically thin carbon films. *science* **306**, 666–669 (2004).

[10] Dean, C. *et al.* Boron nitride substrates for high-quality graphene electronics. *Nature nanotechnology* **5**, 722–726 (2010).

[11] Bistritzer, R. & MacDonald, A. H. Electronic cooling in graphene. *Physical Review Letters* **102**, 206410 (2009).

[12] Tse, W.-K. & Sarma, S. D. Energy relaxation of hot dirac fermions in graphene. *Physical Review B* **79**, 235406 (2009).

[13] Leadley, D., Nicholas, R., Harris, J. & Foxon, C. Influence of acoustic phonons on inter-subband scattering in gaas-gaalas heterojunctions. *Semiconductor Science and Technology* **4**, 885 (1989).

[14] Lutz, J., Kuchar, F., Ismail, K., Nickel, H. & Schlapp, W. Time resolved measurements of the energy relaxation in the 2deg of algaas/gaas. *Semiconductor science and technology* **8**, 399 (1993).

[15] Roukes, M., Freeman, M., Germain, R., Richardson, R. & Ketchen, M. Hot electrons and energy transport in metals at millikelvin temperatures. *Physical review letters* **55**, 422 (1985).

[16] Zuev, Y. M., Chang, W. & Kim, P. Thermoelectric and magnetothermoelectric transport measurements of graphene. *Physical Review Letters* **102**, 096807 (2009).

[17] Wei, P., Bao, W., Pu, Y., Lau, C. N. & Shi, J. Anomalous thermoelectric transport of dirac particles in graphene. *Physical Review Letters* **102**, 166808 (2009).

[18] Nelson, J. *The Physics of Solar Cells* (Imperial College Press, London, 2009).

[19] Lee, E. J. H., Balasubramanian, K., Weitz, R. T., Burghard, M. & Kern, K. Contact and edge effects in graphene devices. *Nature Nanotechnology* **3**, 486 (2008).

[20] Song, J. C. W., Rudner, M. S., Marcus, C. M. & Levitov, L. S. Hot carrier transport and photocurrent response in graphene. *NanoLetters* **11**, 4688 (2011).

[21] Gabor, N. M. & *et. al.* Hot carrier assisted intrinsic photoresponse in graphene. *Science* **334**, 648 (2011).

[22] Yan, J. *et al.* Dual-gated bilayer graphene hot-electron bolometer. *Nature Nanotechnology* **7**, 472 (2012).

[23] Cai, X. *et al.* Sensitive room-temperature terahertz detection via photothermoelectric effect in graphene. *arXiv* 13053297 (2013).

[24] Herring, P. K. *et al.* Sensitive room-temperature terahertz detection via photothermoelectric effect in graphene. *NanoLetters* **14**, 901 (2014).

[25] Rana, F. Electron-hole generation and recombination rates for coulomb scattering in graphene. *Physical Review B* **76**, 155431 (2007).

[26] Song, J. C. W., Reizer, M. Y. & Levitov, L. S. Disorder-assisted electron-phonon scattering and cooling pathways in graphene. *Physical Review Letters* **109**, 106602 (2012).

[27] Graham, M. W., Shi, S.-F., Ralph, D. C., Park, J. W. & McEuen, P. L. Photocurrent measurements of supercollision cooling in graphene. *Nature Physics* **9**, 103 (2013).

[28] Betz, A. C. *et al.* Supercollision cooling in undoped graphene. *Nature Physics* **9**, 109 (2013).

[29] Song, J. C., Tielrooij, K. J., Koppens, F. H. & Levitov, L. S. Photoexcited carrier dynamics and impact-excitation cascade in graphene. *Physical Review B* **87**, 155429 (2013).

[30] Tielrooij, K. *et al.* Photoexcitation cascade and multiple hot-carrier generation in graphene. *Nature Physics* **9**, 248–252 (2013).

[31] Lui, C. H., Mak, K. F. & Heinz, T. F. Ultrafast photoluminescence from graphene. *Physical Review Letters* **105**, 127404 (2010).

[32] Kotov, V. N., Uchoa, B., Pereira, V. M., Guinea, F. & Neto, A. H. C. Electron-electron interactions in graphene: Current status and perspectives. *Reviews of Modern Physics* **84**, 1067 (2012).

[33] Baldo, M. *et al.* Center for excitonics proposal, part 2: Multiexciton and hot carrier dynamics. *EFRC Grant Proposal* (2014).

[34] Ross, R. T. & Nozik, A. J. Efficiency of hot carrier solar energy converters. *Journal of Applied Physics* **52**, 3813 (1982).

[35] Shockley, W. Currents to conductors induced by a moving point charge. *Journal of Applied Physics* **9**, 635 (1938).

[36] Ramo, S. Currents induced by electron motion. *Proceedings of the IRE* **27**, 584 (1939).

[37] Park, J., Ahn, Y. & Ruiz-Vargas, C. Imaging of photocurrent generation and collection in single-layer graphene. *Nano Letters* **9**, 1742 (2009).

[38] Xu, X. *Unpublished* (2012).

[39] Hunt, B. *et al.* Massive dirac fermions and hofstadter butterfly in a van der waals heterostructure. *Science* **340**, 1427 (2013).

[40] Dean, C. R. *et al.* Hofstadter's butterfly and the fractal quantum hall effect in moiré superlattices. *Nature* **497**, 598 (2013).

[41] Ponomarenko, L. A. *et al.* Cloning of dirac fermions in graphene superlattices. *Nature* **497**, 594 (2013).

[42] Park, C.-H., Yang, L., Son, Y.-W., Cohen, M. L. & Louie, S. G. New generation of massless dirac fermions in graphene under external periodic potentials. *Phys. Rev. Lett.* **101**, 126804 (2008).

[43] Park, C.-H., Yang, L., Son, Y.-W., Cohen, M. L. & Louie, S. G. Anisotropic behaviours of massless dirac fermions in graphene under periodic potentials. *Nature Physics* **4**, 213–217 (2008).

[44] Gorbachev, R. V. *et al.* Strong coulomb drag and broken symmetry in double-layer graphene. *Nature Physics* **8**, 896 (2012).

[45] Laikhtman, B. & Solomon, P. M. Mutual drag of two- and three-dimensional electron gases in heterostuctures. *Physical Review B* **41**, 9921 (1990).

[46] Gramila, T. J., Eisenstein, J. P., MacDonald, A. H., Pfeiffer, L. N. & West, K. W. Mutual friction between parallel two-dimensional electron systems. *Physical Review B* **66**, 1216 (1991).

[47] Bid, A. *et al.* Observation of neutral modes in the fractional quantum hall regime. *Nature* **466**, 585–590 (2010).

[48] Venkatachalam, V., Hart, S., Pfeiffer, L., West, K. & Yacoby, A. Local thermometry of neutral modes on the quantum hall edge. *Nature Physics* **8**, 676–681 (2012).

[49] Neto, A. H. C., Guinea, F., Peres, N. M. R., Novosolev, K. & Geim, A. K. The electronic properties of graphene. *Reviews of Modern Physics* **81**, 109 (2009).

[50] Ando, T. Anomaly of optical phonon in monolayer graphene. *Journal of the Physical Society of Japan* **75**, 124701 (2006).

[51] Ma, Q. *et al.* Competing channels for hot electron cooling in graphene. *arXiv preprint arXiv:1403.8152* (2014).

[52] Dawlaty, J. M., Shivaraman, S., Chandrashekhar, M., Rana, F. & Spencer, M. G. Measurement of ultrafast carrier dynamics in epitaxial graphene. *Applied Physics Letters* **92**, 042116 (2008).

[53] Mele, E. J., Král, P. & Tománek, D. Coherent control of photocurrents in graphene and carbon nanotubes. *Physical Review B* **61**, 7669 (2000).

[54] Syzranov, S. V., Fistul, M. V. & Efetov, K. B. Effect of radiation on transport in graphene. *Physical Review B* **78**, 045407 (2008).

[55] Nazin, G., Zhang, Y., Zhang, L., Sutter, E. & Sutter, P. Photocurrent measurements on graphene devices. *Nature Physics* **6**, 870 (2010).

[56] Xu, X., Gabor, N. M., Alden, J. S., van der Zande, A. M. & McEuen, P. L. Photothermoelectric effect at a single-bilayer graphene interface junction. *Nano Letters* **10**, 562 (2010).

[57] Winzer, T., Knorr, A. & Malic, E. Carrier multiplication in graphene. *Nano Letters* **10**, 4839 (2010).

[58] Winzer, T. & Malić, E. Impact of auger processes on carrier dynamics in graphene. *Physical Review B* **85**, 241404 (2012).

[59] Schaller, R. D. & Klimov, V. I. High efficiency carrier multiplication in pbse nanocrystals: Implications for solar energy conversion. *Physical Review Letters* **92**, 186601 (2004).

[60] Califano, M., Zunger, A. & Franceschetti, A. Direct carrier multiplication due to inverse auger scattering in cdse quantum dots. *Applied Physics Letters* **84**, 2409 (2004).

[61] Ellingson, R. J. *et al.* Highly efficient multiple exciton generation in colloidal pbse and pbs quantum dots. *Nano Letters* **5**, 865 (2005).

[62] Nair, G., Chang, L.-Y., Geyer, S. M. & Bawendi, M. G. Perspective on the prospects of a carrier multiplication nanocrystal solar cell. *Nano Letters* **11**, 2145 (2011).

[63] Hwang, E. H., Hu, B. Y. K. & Das Sarma, S. Inelastic carrier lifetime in graphene. *Physical Review B* **76**, 155434 (2007).

[64] Dawlaty, J. H., Shivaraman, S., Chandrashekhar, M., Rana, F. & Spencer, M. G. Measurement of ultrafast carrier dynamics in epitaxial graphene. *Applied Physics Letters* **92**, 042116 (2008).

[65] Efetov, D. & Kim, P. Controlling electron-phonon interactions in graphene at ultrahigh carrier densities. *Physical Review Letters* **105**, 256805 (2010).

[66] Strait, J. H. *et al.* Very slow cooling dynamics of photoexcited carriers in graphene observed by optical-pump terahertz-probe spectroscopy. *NanoLetters* **11**, 4902 (2011).

[67] Lemme, M. C. *et al.* Gate-activated photoresponse in a graphene p-n junction. *Nano Letters* **11**, 4134 (2011).

[68] Hwang, E. H., Rossi, E. & Sarma, S. D. Theory of thermopower in two-dimensional graphene. *Physical Review B* **80**, 235415 (2009).

[69] George, P. A. *et al.* Ultrafast optical-pump terahertz-probe spectroscopy of the carrier relaxation and recombination dynamics in epitaxial graphene. *NanoLetters* **8**, 4248 (2008).

[70] Martin, J. *et al.* Observation of electron-hole puddles in graphene using a scanning single-electron transistor. *Nature Physics* **4**, 144 (2008).

[71] Huard, B. *et al.* Transport measurements across a tunable potential barrier in graphene. *Physical Review Letters* **98**, 236803 (2007).

[72] Williams, J. R., DiCarlo, L. & Marcus, C. M. Quantum hall effect in a gate-controlled pn junction of graphene. *Science* **317**, 638 (2007).

[73] Özyilmaz, B. *et al.* Electronic transport and quantum hall effect in bipolar graphene p-n-p junctions. *Physical Review Letters* **99**, 166804 (2007).

[74] Sun, D. *et al.* Ultrafast hot-carrier-dominated photocurrent in graphene. *Nature nanotechnology* **7**, 114–118 (2012).

[75] Ziman, J. *Principles of the Theory of Solids* (Cambridge University Press, 1972).

[76] Heikkilä, T. T., Luukanen, A., Savin, A. M. & Pekola, J. P. Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications. *Reviews of Modern Physics* **78**, 217 (2006).

[77] Winnerl, S. *et al.* Carrier relaxation in epitaxial graphene photoexcited near the dirac point. *Physical Review Letters* **107**, 237401 (2011).

[78] Ridley, B. K. Hot electrons in low-dimensional structures. *Reports on Progress in Physics* **54**, 169 (1991).

[79] Sarma, S. D., Jain, J. K. & Jalabert, R. Hot electrons in low-dimensional structures. *Physical Review B* **37**, 6290 (1988).

[80] Plochocka, P. *et al.* Slowing hot-carrier relaxation in graphene using a magnetic field. *Physical Review B* **80**, 245415 (2009).

[81] Castro, E. V. *et al.* Limits on charge carrier mobility in suspended graphene due to flexural phonons. *Physical Review Letters* **105**, 266601 (2010).

[82] Ochoa, H., Castro, E. V., Katsnelson, M. I. & Guinea, F. Temperature-dependent resistivity in bilayer graphene due to flexural phonons. *Physical Review B* **83**, 235416 (2011).

[83] Fratini, S. & Guinea, F. Substrate-limited electron dynamics in graphene. *Physical Review B* **77**, 195415 (2008).

[84] Price, A. S., Hornett, S. M., Shytov, A. V., Hendry, E. & Horsell, D. W. Nonlinear resistivity and heat dissipation in monolayer graphene. *Physical Review B* **85**, 161411 (2012).

[85] Morpurgo, A. F. & Guinea, F. Intervalley scattering, long-range disorder, and effective time-reversal symmetry breaking in graphene. *Physical Review Letters* **97**, 196804 (2006).

[86] Katsnelson, M. I. & Geim, A. Electron scattering on microscopic corrugations in graphene. *Philosophical Transactions of the Royal Society A* **366**, 195 (2008).

[87] Kim, E.-A. & Neto, A. C. Graphene as an electronic membrane. *EuroPhysics Letters* **84**, 57007 (2008).

[88] Vozmediano, M., Katsnelson, M. & Guinea, F. Gauge fields in graphene. *Physics Reports* **496**, 109 (2010).

[89] Ishigami, M., Chen, J. H., Cullen, W. G., Fuhrer, M. S. & Williams, E. D. Atomic structure of graphene on sio2. *NanoLetters* **9**, 1644 (2007).

[90] Lundeberg, M. & Folk, J. Rippled graphene in an in-plane magnetic field: Effects of a random vector potential. *Physical Review Letters* **105**, 146804 (2010).

[91] Shon, N. H. & Ando, T. Quantum transport in two-dimensional graphite system. *Journal of the Physical Society of Japan* **67**, 2421 (1998).

[92] Viljas, J. K. & Heikkilä, T. T. Electron-phonon heat transfer in monolayer and bilayer graphene. *Physical Review B* **81**, 245404 (2010).

[93] Pisana, S. & *et. al.* Breakdown of the adiabatic born-oppenheimer approximation in graphene. *Physical Review Letters* **6**, 198 (2007).

[94] Bonini, N., Lazerri, M., Marzari, N. & Mauri, F. Phonon anharmonicities in graphite and graphene. *Physical Review Letters* **99**, 176802 (2007).

[95] Graham, M. W., Shi, S.-F., Ralph, D. C., Park, J. & McEuen, P. L. Photocurrent measurements of supercollision cooling in graphene. *Physical Review B* **9**, 103 (2012).

[96] Xu, X. *Private Communication* (2012).

[97] Kampfrath, T., Perfetti, L., Schapper, F., Frischkorn, C. & Wolf, M. Strongly coupled optical phonons in the ultrafast dynamics of the electronic energy and current relaxation in graphite. *Physical review letters* **95**, 187403 (2005).

[98] Foster, M. S. & Aleiner, I. L. Slow imbalance relaxation and thermoelectric transport in graphene. *Physical Review B* **79**, 085415 (2009).

[99] Butscher, S., Milde, F., Hirtschulz, M., Malić, E. & Knorr, A. Hot electron relaxation and phonon dynamics in graphene. *Applied Physics Letters* **91**, 203103 (2007).

[100] Vasko, F. & Ryzhii, V. Photoconductivity of intrinsic graphene. *Physical Review B* **77**, 195433 (2008).

[101] Kim, R., Perebeinos, V. & Avouris, P. Relaxation of optically excited carriers in graphene. *Physical Review B* **84**, 075449 (2011).

[102] Breusing, M., Ropers, C. & Elsaesser, T. Ultrafast carrier dynamics in graphite. *Physical review letters* **102**, 086809 (2009).

[103] Gierz, I. *et al.* Snapshots of non-equilibrium dirac carrier distributions in graphene. *Nature materials* (2013).

[104] Gonzales, J., Guinea, F. & Vozmediano, M. Non-fermi liquid behavior of electrons in the half-filled honeycomb lattice (a renormalization group approach). *Nuclear Physics B* **424**, 595 (1994).

[105] Elias, D. *et al.* Dirac cones reshaped by interaction effects in suspended graphene. *Nature Physics* **7**, 701–704 (2011).

[106] Zhou, S. *et al.* First direct observation of dirac fermions in graphite. *Nature physics* **2**, 595–599 (2006).

[107] Bostwick, A., Ohta, T., Seyller, T., Horn, K. & Rotenberg, E. Quasiparticle dynamics in graphene. *Nature Physics* **3**, 36–40 (2007).

[108] Polini, M. *Private Communication* (2013).

[109] Basko, D., Piscanec, S. & Ferrari, A. Electron-electron interactions and doping dependence of the two-phonon raman intensity in graphene. *Physical Review B* **80**, 165413 (2009).

[110] Polini, M. *et al.* Plasmons and the spectral function of graphene. *Physical Review B* **77**, 081411 (2008).

[111] Hwang, E. & Sarma, S. D. Dielectric function, screening, and plasmons in two-dimensional graphene. *Physical Review B* **75**, 205418 (2007).

[112] Giuliani, G. F. & Quinn, J. J. Lifetime of a quasiparticle in a two-dimensional electron gas. *Physical Review B* **26**, 4421 (1982).

[113] Jackson, J. D. *Classical electrodynamics*, vol. 1 (Wiley, 1998).

[114] Beard, M. C., Turner, G. M. & Schmuttenmaer, C. A. Transient photoconductivity in gaas as measured by time-resolved terahertz spectroscopy. *Physical Review B* **62**, 15764 (2000).

[115] Frenzel, A. *et al.* Observation of suppressed terahertz absorption in photoexcited graphene. *Applied Physics Letters* **102**, 113111 (2013).

[116] Jnawali, G., Rao, Y., Yan, H. & Heinz, T. F. Observation of a transient decrease in terahertz conductivity of single-layer graphene induced by ultrafast optical excitation. *Nano letters* **13**, 524–530 (2013).

[117] Shi, S.-F. *et al.* Controlling graphene ultrafast hot carrier response from metal-like to semiconductor-like by electrostatic gating. *Nano letters* **14**, 1578–1582 (2014).

[118] Nomura, K. & MacDonald, A. H. Quantum hall ferromagnetism in graphene. *Physical review letters* **96**, 256602 (2006).

[119] Sarma, S. D., Adam, S., Hwang, E. H. & Rossi, E. Electronic transport in two-dimensional graphene. *Reviews of Modern Physics* **83**, 407 (2011).

[120] Nieva, J. R. & Song, J. C. W. Hot carrier extraction in graphene heterostructures. *Manuscript in preparation* (2014).

[121] Sze, S. M. & Ng, K. K. *Physics of Semiconductor Devices, 3rd Edition* (Wiley, New York, 2007).

[122] Deyo, E., Golub, L. E., Ivchenko, E. L. & Spivak, B. Semiclassical theory of the photogalvanic effect in non-centrosymmetric systems. *arXiv* 0904.1917 (2009).

[123] Moore, J. & Orenstein, J. Confinement-induced berry phase and helicity-dependent photocurrents. *Physical Review Letters* **105**, 026805 (2010).

[124] Hosur, P. Circular photogalvanic effect on topological insulator surfaces: Berry-curvature-dependent response. *Physical Review B* **83**, 035309 (2010).

[125] Rana, F. *et al.* Carrier recombination and generation rates for intravalley and intervalley phonon scattering in graphene. *Physical Review B* **79**, 115447 (2011).

[126] He, Z. Review of the shockleyÐramo theorem and its application in semiconductor gamma-ray detectors. *Nuclear Institute Methods of Physics Research A* **250**, 463 (2001).

[127] Xia, F. *et al.* Photocurrent imaging and efficient photon detection in a graphene transistor. *Nano Letters* **9**, 1039 (2009).

[128] McIver, J. W., Hsieh, D., Steinberg, H., P., J.-H. & Gedik, N. Control over topological insulator photocurrents with light polarization. *Nature Nanotechnology* **7**, 96 (2011).

[129] Yoder, P., Gärtner, K. & Fichtner, W. A generalized ramo-shockley theorem for classical to quantum transport at arbitrary frequencies. *Journal of Applied Physics* **79**, 1952 (1996).

[130] Maero, G., Paroli, B., Pozzoli, R. & Romé, M. Stabilizing effect of a nonresonant radio frequency drive on the m=1 diocotron instability. *Physics of Plasmas* **18**, 032101 (2011).

[131] Cavalleri, G., Gatti, E., Fabri, G. & Svelto, V. Extension of ramo's theorem as applied to induced charge in semiconductor detectors. *Nuclear Instrumentation and Methods* **92**, 137 (1971).

[132] Chklovskii, D. B., Shklovskii, B. & Glazman, L. Electrostatics of edge channels. *Physical Review B* **83**, 035309 (1992).

[133] Jamei, R., Kivelson, S. & Spivak, B. Universal aspects of coulomb-frustrated phase separation. *Physical Review Letters* **94**, 056805 (2005).

[134] Britnell, L. *et al.* Field-effect tunneling transistor based on vertical graphene heterostructures. *Science* **335**, 947 (2012).

[135] Tse, W.-K. & Sarma, S. D. Coulomb drag and spin drag in the presence of spin-orbit coupling. *Physical Review B* **75**, 045333 (2007).

[136] Hwang, E. H., Sensarma, R. & Sarma, S. D. Coulomb drag in monolayer and bilayer graphene. *Physical Review B* **84**, 245441 (2011).

[137] Narozhny, B. N. Coulomb drag as a measure of trigonal warping in doped graphene. *Physical Review B* **76**, 153409 (2007).

[138] Peres, N. M. R., dos Santos, J. M. B. L. & Neto, A. H. C. Coulomb drag and high-resistivity behavior in double-layer graphene. *Europhysics Letters* **95**, 18001 (2011).

[139] Katsnelson, M. I. Coulomb drag in graphene single layers separated by a thin spacer. *Physical Review B* **84**, 041407 (2011).

[140] Narozhny, B. N., Titov, M., Gornyi, I. V. & Ostrovsky, P. M. Coulomb drag in graphene: Perturbation theory. *Physical Review B* **85**, 195421 (2012).

[141] Kim, S. *et al.* Coulomb drag of massless fermions in graphene. *Physical Review B* **83**, 161401 (2011).

[142] Sivan, U., Solomon, P. M. & Shtrikman, H. Coupled electron-hole transport. *Physical Review Letters* **68**, 1196 (1992).

[143] Zhang, Y., Brar, V. W., Girit, C., Zettl, A. & Crommie, M. F. Origin of spatial charge inhomogeneity in graphene. *Nature Physics* **6**, 722 (2009).

[144] Decker, R. *et al.* Local electronic properties of graphene on a bn substrate via scanning tunneling microscopy. *NanoLetters* **11**, 4631 (2011).

[145] Chen, J.-H., Jang, C., Xiao, S., Ishigami, M. & Fuhrer, M. S. Intrinsic and extrinsic performance limits of graphene devices on sio2. *Nature Nanotechnology* **3**, 206 (2008).

[146] Andreev, A. V., Kivelson, S. A. & Spivak, B. Hydrodynamic description of transport in strongly correlated electron systems. *Physical Review Letters* **106**, 256804 (2011).

[147] Ando, T. Screening effect and impurity scattering in monolayer graphene. *Journal of the Physical Society of Japan* **75**, 074716 (2006).

[148] Gibertini, M., Tomadin, A., Guinea, F., Katsnelson, M. I. & Polini, M. Electron-hole puddles in the absence of charged impurities. *Physical Review B* **85**, 201405 (2012).

[149] Sarma, S. D., Hwang, E. H. & Tse, W.-K. Many-body interaction effects in doped and undoped graphene: Fermi liquid versus non-fermi liquid. *Physical Review B* **75**, 121406 (2007).

[150] Müller, M., Fritz, L. & Sachdev, S. Quantum-critical relativistic magnetotransport in graphene. *Physical Review B* **78**, 115406 (2008).

[151] Lyakhov, A. O. & Mishchenko, E. G. Thermal conductivity of a two-dimensional electron gas with coulomb interaction. *Physical Review B* **67**, 041304 (2003).

[152] Callen, H. B. The application of onsager's reciprocal relations to thermoelectric, thermomagnetic, and galvanomagnetic effects. *Physical Review* **73**, 1349 (1948).

[153] Gonzales, J., Guinea, F. & Vozmediano, M. Marginal-fermi-liquid behavior from two-dimensional coulomb interaction. *Physical Review B* **59**, R2474 (1999).

[154] Sheehy, D. E. & Schmalian, J. Quantum critical scaling in graphene. *Physical Review Letters* **99**, 226803 (2007).

[155] Son, D. T. Quantum critical point in graphene approached in the limit of infinitely strong coulomb interaction. *Physical Review B* **75**, 235423 (2007).

[156] Vafek, O. Anomalous thermodynamics of coulomb-interacting massless dirac fermions in two spatial dimensions. *Physical Review Letters* **98**, 216401 (2007).

[157] Kashuba, A. B. Conductivity of defectless graphene. *Physical Review B* **78**, 085415 (2008).

[158] Fritz, L., Schmalian, J., Müller, M. & Sachdev, S. Quantum critical transport in clean graphene. *Physical Review B* **78**, 085416 (2008).

[159] Checkelsky, J. G. & Ong, N. P. Thermopower and nernst effect in graphene in a magnetic field. *Physical Review B* **80**, 081413 (2009).

[160] Pitaevskii, L. P. & Lifshitz, E. *Physical Kinetics* (Pergamon, Oxford, 1981).

[161] Hwang, H. Y. *et al.* Nonlinear thz conductivity dynamics in cvd-grown graphene. *arXiv preprint arXiv:1101.4985* (2011).

[162] Song, J. C. W. & Levitov, L. S. Energy-driven drag at charge neutrality in graphene. *Physical Review Letters* **109**, 236602 (2012).

[163] Song, J. C. W., Abanin, D. A. & Levitov, L. S. The mechanisms of coulomb drag. *Nano Letters* **109**, 236602 (2013).

[164] Patel, N. K. *et al.* Magnetic field studies of coulomb drag in a coupled double quantum well system. *Semiconductor Science and Technology* **12**, 309 (1997).

[165] Rubel, H., Fischer, A., Dietsche, W., von Klitzing, K. & Eberl, K. Observation of screening in the magneto-coulomb drag between coupled two-dimensional electron systems. *Physical Review Letters* **78**, 1763 (1997).

[166] Lilly, M. P., Eisenstein, J. P., Pfeiffer, L. N. & West, K. W. Coulomb drag in the extreme quantum limit. *Physical Review Letters* **80**, 1714 (1998).

[167] Kamenev, A. & Oreg, Y. Coulomb drag in normal metals and superconductors: Diagrammatic approach. *Physical Review B* **52**, 7516 (1995).

[168] Bonsager, M. C., Flensberg, K., Hu, B. Y.-K. & Jauho, A.-P. Magneto-coulomb drag: Interplay of electron-electron interactions and landau quantization. *Physical Review Letters* **77**, 1366 (1996).

[169] Jonson, M. & Girvin, S. M. Thermoelectric effect in a weakly disordered inversion layer subject to a quantizing magnetic field. *Physical Review B* **29**, 1939 (1984).

[170] Abanin, D. A. *et al.* Giant nonlocality near the dirac point in graphene. *Science* **332**, 328–330 (2011).

[171] Geim, A. K. *Private Communication* (2013).

[172] Jauho, A.-P. & Smith, H. Coulomb drag between parallel two-dimensional electron systems. *Physical Review B* **47**, 4420 (1993).

[173] Zheng, L. & MacDonald, A. H. Coulomb drag between disordered two-dimensional electron-gas layers. *Physical Review B* **48**, 8203 (1993).

[174] Flensberg, K., Hu, B. Y.-K. & Jauho, A.-P. Linear-response theory of coulomb drag in coupled electron systems. *Physical Review B* **52**, 14761 (1995).

[175] Song, J. C. W. & Levitov, L. S. Hall drag and magnetodrag in graphene. *Physical Review Letters* **111**, 126601 (2013).

[176] Gantmakher, V. F. & Levinson, Y. B. Effect of collisions between carriers of the dissipative carriers. *Soviet Physics, Journal of Experimental and Theoretical Physics* **47**, 133 (1978).

[177] Abanin, D. A., Gorbachev, R. V., Novoselov, K. S., Geim, A. K. & Levitov, L. S. Giant spin-hall effect induced by the zeeman interaction in graphene. *Physical Review Letters* **107**, 096601 (2011).

[178] Neto, A. C. & Guinea, F. Impurity-induced spin-orbit coupling in graphene. *Physical review letters* **103**, 026804 (2009).

[179] Smith, H. & Jensen, H. H. *Transport phenomena, Sec 4.3: Second Sound* (Oxford University Press, USA, 1989).

[180] Peshkov, V. Second sound in helium ii. *J. Phys. USSR* **8**, 381 (1944).

[181] Narayanamurti, V. & Dynes, R. Observation of second sound in bismuth. *Physical Review Letters* **28**, 1461 (1972).

[182] Landau, L. & Lifshitz, E. Fluid mechanics, vol. 6. *Course of Theoretical Physics* 227–229 (1987).

[183] Sunyaev, R. A. & Zeldovich, Y. B. Small-scale fluctuations of relic radiation. *Astrophysics and Space Science* **7**, 3–19 (1970).

[184] Peebles, P. J. & Yu, J. Primeval adiabatic perturbation in an expanding universe. *The Astrophysical Journal* **162**, 815 (1970).

[185] Chen, J. *et al.* Optical nano-imaging of gate-tunable graphene plasmons. *Nature* (2012).

[186] Fei, Z. *et al.* Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature* (2012).

[187] Bolotin, K. I. *et al.* Ultrahigh electron mobility in suspended graphene. *Solid State Communications* **146**, 351–355 (2008).

[188] Taychatanapat, T., Watanabe, K., Taniguchi, T. & Jarillo-Herrero, P. Electrically tunable transverse magnetic focusing in graphene. *Nature Physics* **9**, 225–229 (2013).

[189] Müller, M., Schmalian, J. & Fritz, L. Graphene: A nearly perfect fluid. *Physical review letters* **103**, 025301 (2009).

[190] Vafek, O. Thermoplasma polariton within scaling theory of single-layer graphene. *Physical review letters* **97**, 266406 (2006).

[191] Gangadharaiah, S., Farid, A. & Mishchenko, E. Charge response function and a novel plasmon mode in graphene. *Physical review letters* **100**, 166802 (2008).

[192] Sarma, S. D. & Li, Q. Intrinsic plasmons in two-dimensional dirac materials. *Physical Review B* **87**, 235418 (2013).

[193] Brida, D. *et al.* Ultrafast collinear scattering and carrier multiplication in graphene. *Nature communications* **4** (2013).